Access DB# 135 465

SEARCH REQUEST FORM
Scientific and Technical Information Center
(STIC)

CRFE

OCT 28 2004

Requester's Full Name: __Jane ZARA__ Examiner #: __77518__ Date: __10/14/04__

Art Unit: __1635__ Phone Number __30 2-0765__ Serial Number: __09/471 067__

Mail Box and Bldg/Room Location: __2d 28__ Results Format Preferred (circle): **PAPER** DISK E-MAIL
2C18

If more than one search is submitted, please prioritize searches in order of need. __ME__

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: __NA Ladders__

Inventors (please provide full names): __Hu et al.__

Earliest Priority Filing Date: __3 - 26 - 98__

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search Seq IDs 1 - 4. (34) (seq 067)

— in the following arrangement:

ATCTCAGGAT   ATCCTGAGAT

NA-10-(1) seg 1 + 1   or seg 1 + complement of 1

ATCAGTCGAT   ATCGACTGAT

NA-10-(2) seg 2 + 2   or seg 2 + comple of 2

ATCGCATGAT   ATCATGCGAT

NA-10-(3) Seg 3 + 3   or seg 3 + comple of 3

ATCATGCGAT   ATCGCATGAT

NA-10 (4) Seg 4 + 4   or seg 4 + comple of 4

ME'

For interference & regular data bases please.

EXACT IDentity needed.

****

PTO-1590 (8-01)

File 159:Cancerlit 1975-2002/Oct
        (c) format only 2002 Dialog Corporation
File 164:Allied & Complementary Medicine 1984-2004/Dec
        (c) 2004 BLHCIS
File 444:New England Journal of Med. 1985-2004/Nov W4
        (c) 2004 Mass. Med. Soc.
File 467:ExtraMED(tm) 2000/Dec
        (c) 2001 Informania Ltd.

Set   Items   Description
S1   0   NUCLEIC ACID LADDER?
S2   560   DNA LADDER?
S3   0   S2 (S) RESTRICTION
S4   27   S2 AND ENDONUCLEASE?
S5   25   RD (unique items)
>>>KWIC option is not available in file(s): 399

 5/3,K/1   (Item 1 from file: 5)
DIALOG(R)File   5:Biosis Previews(R)
(c) 2004 BIOSIS. All rts. reserv.

0014837034   BIOSIS NO.: 200400204667
**Mechanism of enhanced vulnerability to NMDA toxicity in secondary traumatic
    neuronal injury.**
AUTHOR: Arundine M (Reprint); Chopra G K; Wrong A (Reprint); Lei S
    (Reprint); Aarts M M; Wallace M; Macdonald J F (Reprint); Tymianski M
    (Reprint)
AUTHOR ADDRESS: Physiology, Univ. of Toronto, Toronto Western Res. Inst.,
    Toronto, ON, Canada**Canada
JOURNAL: Society for Neuroscience Abstract Viewer and Itinerary Planner
2003 pAbstract No. 743.2 2003 2003
MEDIUM: e-file
CONFERENCE/MEETING: 33rd Annual Meeting of the Society of Neuroscience New
Orleans, LA, USA November 08-12, 2003; 20031108
SPONSOR: Society of Neuroscience
DOCUMENT TYPE: Meeting; Meeting Abstract
RECORD TYPE: Abstract
LANGUAGE: English

...ABSTRACT: a pan caspase inhibitor z-vad-FMK did not attenuate
    vulnerability to NMDA, suggesting against a classical apoptotic
    mechanism. Moreover, apoptosis-inducing factor (AIF) and *endonuclease*-G
    did not translocate into the nucleus, suggesting against caspase
    independent apoptotic mechanisms. Stretched neurons exhibited a rapid
    production of reactive oxygen and nitrogen species...
DESCRIPTORS:
    METHODS & EQUIPMENT: *DNA laddering*...

 5/3,K/2   (Item 2 from file: 5)
DIALOG(R)File   5:Biosis Previews(R)
(c) 2004 BIOSIS. All rts. reserv.

0014301368   BIOSIS NO.: 200300260012
**Effect of hypoxia on DNA fragmentation in the cerebral cortex of the guinea
    pig fetus at different gestational ages.**
AUTHOR: Abedin Naheed (Reprint); Ashraf Oazi M (Reprint); Randis Tara
    (Reprint); Mishra Om P (Reprint); Delivoria-Papadopoulos Maria (Reprint)
AUTHOR ADDRESS: Dept. of Pediatrics, Drexel University and St.
    Christopher's Hospital for Children, Philadelphia, PA, USA**USA
JOURNAL: Pediatric Research 53 (4 Part 2): p48A-49A April 2003 2003
MEDIUM: print
CONFERENCE/MEETING: Annual Meeting of the Pediatric Academic Societies'
Seattle, WA, USA May 03-06, 2003; 20030503
SPONSOR: Pediatric Academic Societies
ISSN: 0031-3998 _(ISSN print)
DOCUMENT TYPE: Meeting; Meeting Abstract

```
File    5:Biosis Previews(R) 1969-2004/Nov W3
          (c) 2004 BIOSIS
File    6:NTIS 1964-2004/Nov W4
          (c) 2004 NTIS, Intl Cpyrght All Rights Res
File    8:Ei Compendex(R) 1970-2004/Nov W4
          (c) 2004 Elsevier Eng.  Info. Inc.
File   34:SciSearch(R) Cited Ref Sci 1990-2004/Nov W4
          (c) 2004 Inst for Sci Info
File   65:Inside Conferences 1993-2004/Dec W1
          (c) 2004 BLDSC all rts. reserv.
File   71:ELSEVIER BIOBASE 1994-2004/Nov W4
          (c) 2004  Elsevier Science B.V.
File   73:EMBASE 1974-2004/Nov W4
          (c) 2004 Elsevier Science B.V.
File   94:JICST-EPlus 1985-2004/Oct W5
          (c)2004 Japan Science and Tech Corp(JST)
File   98:General Sci Abs/Full-Text 1984-2004/Sep
          (c) 2004 The HW Wilson Co.
File   99:Wilson Appl. Sci & Tech Abs 1983-2004/Oct
          (c) 2004 The HW Wilson Co.
File  135:NewsRx Weekly Reports 1995-2004/Nov W4
          (c) 2004 NewsRx
File  143:Biol. & Agric. Index 1983-2004/Oct
          (c) 2004 The HW Wilson Co
File  144:Pascal 1973-2004/Nov W4
          (c) 2004 INIST/CNRS
File  155:MEDLINE(R) 1951-2004/Nov W4
          (c) format only 2004 The Dialog Corp.
File  172:EMBASE Alert 2004/Nov W4
          (c) 2004 Elsevier Science B.V.
File  266:FEDRIP 2004/Sep
          Comp & dist by NTIS, Intl Copyright All Rights Res
File  315:ChemEng & Biotec Abs 1970-2004/Nov
          (c) 2004 DECHEMA
File  357:Derwent Biotech Res. _1982-2004/Dec W2
          (c) 2004 Thomson Derwent & ISI
File  358:Current BioTech Abs 1983-2004/Nov
          (c) 2004 DECHEMA
File  369:New Scientist 1994-2004/Nov W4
          (c) 2004 Reed Business Information Ltd.
File  370:Science 1996-1999/Jul W3
          (c) 1999 AAAS
File  399:CA SEARCH(R) 1967-2004/UD=14124
          (c) 2004 American Chemical Society
File  434:SciSearch(R) Cited Ref Sci 1974-1989/Dec
          (c) 1998 Inst for Sci Info
File   40:Enviroline(R) 1975-2004/Oct
File   50:CAB Abstracts 1972-2004/Oct
          (c) 2004 CAB International
File  103:Energy SciTec 1974-2004/Nov B2
          (c) 2004 Contains copyrighted material
File  156:ToxFile 1965-2004/Nov W2
          (c) format only 2004 The Dialog Corporation
File  162:Global Health 1983-2004/Oct
          (c) 2004 CAB International
File  305:Analytical Abstracts 1980-2004/Dec W1
          (c) 2004 Royal Soc Chemistry
File  393:Beilstein Abstracts  July 2004
          (c) 2004 Beilstein GmbH
File   35:Dissertation Abs Online 1861-2004/Nov
          (c) 2004 ProQuest Info&Learning
File   48:SPORTDiscus 1962-2004/Dec
          (c) 2004 Sport Information Resource Centre
File   91:MANTIS(TM) 1880-2004/Nov
          2001 (c) Action Potential
File  149:TGG Health&Wellness DB(SM) 1976-2004/Nov W1
          (c) 2004 The Gale Group
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.　　GenCore version 5.1.6

Run on:　　October 27, 2004, 11:49:55 ; Search time 710.125 Seconds
　　　　　　(without alignments)
　　　　　　1331.870 Million cell updates/sec

Title:　　　　　SEQ4-C4
Perfect score:　20
Sequence:　　　1 atcacggatatcgcatgat 20
Scoring table:　OLIGO_NUC
　　　　　　　　Gapop 60.0 , Gapext 60.0
Searched:　　　4526729 seqs, 2364849745 residues
Word size :　　0

Total number of hits satisfying chosen parameters:　9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
　　　1: gb_ba:*
　　　　GenEmbl:*
　　　2: gb_htg:*
　　　3: gb_in:*
　　　4: gb_om:*
　　　5: gb_ov:*
　　　6: gb_pat:*
　　　7: gb_ph:*
　　　8: gb_pl:*
　　　9: gb_p:*
　　10: gb_ro:*
　　11: gb_sts:*
　　12: gb_sy:*
　　13: gb_un:*
　　14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match Length	DB	ID	Description
1	16	80.0	427	1	AJ621974	AJ621974 Unculture
2	16	80.0	427	1	AJ621974	AJ621974 Unculture
3	16	80.0	429	1	AJ621971	AJ621971 Unculture
4	16	80.0	429	1	AJ621971	AJ621971 Unculture
5	16	80.0	745	1	AY102896	AY102896 Unculture
6	16	80.0	745	1	AY102896	AY102896 Unculture
7	16	80.0	814	1	AJ387881	AJ387881 Unidentif
8	16	80.0	814	1	UBA387881	UBA387881 Unidentif
9	16	80.0	19358	2	AC136479	AC136479 Mus muscu
10	16	80.0	19358	2	AC136479	AC136479 Mus muscu
11	16	80.0	210346	2	BX511154	BX511154 Danio rer
12	16	80.0	210346	2	BX511154	BX511154 Danio rer
13	16	80.0	244788	5	BX511023	BX511023 Zebrafish
14	16	80.0	244788	5	BX511023	BX511023 Zebrafish
15	15	75.0	504	1	AF261058	AF261058 Gelidibac
16	15	75.0	504	1	AF261058	AF261058 Gelidibac
17	15	75.0	1394	1	AY162084	AY162084 Bacteroid
18	15	75.0	1394	1	AY162084	AY162084 Bacteroid
19	15	75.0	1426	1	AF218782	AF218782 Muricauda
20	15	75.0	1426	1	AF218782	AF218782 Muricauda
21	15	75.0	1491	1	AF300973	AF300973 Bacterium
22	15	75.0	1491	1	AF300973	AF300973 Bacterium
23	15	75.0	190588	2	AC141720	AC141720 Apis mell
24	15	75.0	190588	2	AC141720	AC141720 Apis mell
25	15	75.0	191208	2	AC138026	AC138026 Mus muscu
26	15	75.0	191208	2	AC138026	AC138026 Mus muscu
27	15	75.0	221150	2	AC134589	AC134589 Mus muscu
28	15	75.0	221150	2	AC134589	AC134589 Mus muscu
29	15	75.0	300099	1	AE016764	AE016764 Escherich
30	15	75.0	300099	1	AE016764	AE016764 Escherich
31	15	75.0	349723	1	BX842650	BX842650 Bdellovib
32	15	75.0	349723	1	BX842650	BX842650 Bdellovib
33	14	70.0	376	1	AY006733	AY006733 Unculture
34	14	70.0	376	1	AY006733	AY006733 Unculture
35	14	70.0	499	1	AB028348	AB028348 Unidentif
36	14	70.0	499	1	AB028348	AB028348 Unidentif
37	14	70.0	499	1	AB028358	AB028358 Unidentif
38	14	70.0	499	1	AB028358	AB028358 Unidentif
39	14	70.0	503	1	AB121810	AB121810 Streptoco
40	14	70.0	503	1	AB121810	AB121810 Streptoco
41	14	70.0	517	1	AB028427	AB028427 Unidentif
42	14	70.0	517	1	AB028427	AB028427 Unidentif
43	14	70.0	529	1	AB064836	AB064836 Unculture
44	14	70.0	529	1	AB064836	AB064836 Unculture
45	14	70.0	531	1	AB064835	AB064835 Unculture

ALIGNMENTS

RESULT 1
AJ621974

LOCUS　　　　AJ621974　　　　　427 bp　　　DNA　　　linear　　BCT 23-JUL-2004
DEFINITION　Uncultured bacterium partial 16S rRNA gene, clone IRR-DS7-20.
ACCESSION　AJ621974
VERSION　　AJ621974.1　GI:50582380
KEYWORDS　16S ribosomal RNA; 16S rRNA gene.
SOURCE　　　uncultured bacterium
ORGANISM　uncultured bacterium
　　　　　　　Bacteria; environmental samples.
REFERENCE　1
AUTHORS　　Scheid,D., Stubner,S. and Conrad,R.
TITLE　　　Identification of rice root associated nitrate, sulfate and ferric
　　　　　　　iron reducing bacteria
JOURNAL　　Unpublished
REFERENCE　2　(bases 1 to 427)
AUTHORS　　Scheid,D.
TITLE　　　Direct Submission
JOURNAL　　Submitted (23-JAN-2004) Scheid D., Biogeochemistry, MPI for
　　　　　　　Terrestrial Microbiology, Karl-Von-Frisch Str., D-35043 Marburg,
　　　　　　　GERMANY

FEATURES　　　　　　　　Location/Qualifiers
　　source　　　　　1..427
　　　　　　　　　　　/organism="uncultured bacterium"
　　　　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　　　　/isolation_source="roots"
　　　　　　　　　　　/specific_host="Oryza sativa"
　　　　　　　　　　　/db_xref="taxon:77133"
　　　　　　　　　　　/clone="IRR-DS7-20"
　　　　　　　　　　　/country="Germany"
　　gene　　　　　　1..427
　　　　　　　　　　　/gene="16S rRNA"
　　rRNA　　　　　1..427
　　　　　　　　　　　/gene="16S rRNA"
　　　　　　　　　　　/product="16S ribosomal RNA"

ORIGIN

Query Match　　80.0%;　Score 16;　DB 1;　Length 427;
Best Local Similarity　100.0%;　Pred. No. 52;
Matches　16;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　　5 TGCGGATATCGCATGAT 20

Db   144 TGCGATATCGGATGAT 159
      ||||||||||||||||
Qy   1 TGCGATATCGGATGAT 16

RESULT 2
LOCUS   AJ621974/c
DEFINITION   Uncultured bacterium partial 16S rRNA gene, clone IRR-DS7-20.
ACCESSION   AJ621974
VERSION   AJ621974.1 GI:50582380
KEYWORDS   16S ribosomal RNA; 16S rRNA gene.
SOURCE   uncultured bacterium
  ORGANISM   uncultured bacterium
             Bacteria; environmental samples.
REFERENCE   1
  AUTHORS   Scheid,D., Stubner,S. and Conrad,R.
  TITLE   Identification of rice root associated nitrate, sulfate and ferric
             iron reducing bacteria
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 427)
  AUTHORS   Scheid,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-JAN-2004) Scheid D., Biogeochemistry, MPI for
             Terrestrial Microbiology, Karl-Von-Frisch Str., D-35043 Marburg,
             GERMANY

FEATURES   Location/Qualifiers
     source   1..427
                     /organism="uncultured bacterium"
                     /mol_type="genomic DNA"
                     /isolation_source="roots"
                     /specific_host="Oryza sativa"
                     /db_xref="taxon:77133"
                     /clone="IRR-DS7-20"
                     /country="Germany"
     gene   1..427
                     /gene="16S rRNA"
     rRNA   1..427
                     /gene="16S rRNA"
                     /product="16S ribosomal RNA"

ORIGIN

Query Match   80.0%; Score 16; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 52;
Matches   16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATCATGCGATATCGCA 16
          ||||||||||||||||
Db   159 ATCATGCGGATATCGCA 144

RESULT 3
LOCUS   AJ621971   429 bp   DNA   linear   BCT 23-JUL-2004
DEFINITION   Uncultured bacterium partial 16S rRNA gene, clone IRR-DS7-8.
ACCESSION   AJ621971
VERSION   AJ621971.1 GI:50582377
KEYWORDS   16S ribosomal RNA; 16S rRNA gene.
SOURCE   uncultured bacterium
  ORGANISM   uncultured bacterium
             Bacteria; environmental samples.
REFERENCE   1
  AUTHORS   Scheid,D., Stubner,S. and Conrad,R.
  TITLE   Identification of rice root associated nitrate, sulfate and ferric
             iron reducing bacteria
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 429)
  AUTHORS   Scheid,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-JAN-2004) Scheid D., Biogeochemistry, MPI for
             Terrestrial Microbiology, Karl-Von-Frisch Str., D-35043 Marburg,
             GERMANY

FEATURES   Location/Qualifiers
     source   1..429
                     /organism="uncultured bacterium"
                     /mol_type="genomic DNA"
                     /isolation_source="roots"
                     /specific_host="Oryza sativa"
                     /db_xref="taxon:77133"
                     /clone="IRR-DS7-8"
                     /country="Germany"
     gene   1..429
                     /gene="16S rRNA"
     rRNA   1..429
                     /gene="16S rRNA"
                     /product="16S ribosomal RNA"

ORIGIN

Query Match   80.0%; Score 16; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 52;
Matches   16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   5 TGCGATATCGGATGAT 20
          ||||||||||||||||
Db   144 TGCGATATCGGATGAT 159

RESULT 4
LOCUS   AJ621971/c
DEFINITION   Uncultured bacterium partial 16S rRNA gene, clone IRR-DS7-8.
ACCESSION   AJ621971
VERSION   AJ621971.1 GI:50582377
KEYWORDS   16S ribosomal RNA; 16S rRNA gene.
SOURCE   uncultured bacterium
  ORGANISM   uncultured bacterium
             Bacteria; environmental samples.
REFERENCE   1
  AUTHORS   Scheid,D., Stubner,S. and Conrad,R.
  TITLE   Identification of rice root associated nitrate, sulfate and ferric
             iron reducing bacteria
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 429)
  AUTHORS   Scheid,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-JAN-2004) Scheid D., Biogeochemistry, MPI for
             Terrestrial Microbiology, Karl-Von-Frisch Str., D-35043 Marburg,
             GERMANY

FEATURES   Location/Qualifiers
     source   1..429
                     /organism="uncultured bacterium"
                     /mol_type="genomic DNA"
                     /isolation_source="roots"
                     /specific_host="Oryza sativa"
                     /db_xref="taxon:77133"
                     /clone="IRR-DS7-8"
                     /country="Germany"
     gene   1..429
                     /gene="16S rRNA"
     rRNA   1..429
                     /gene="16S rRNA"
                     /product="16S ribosomal RNA"

ORIGIN

Query Match   80.0%; Score 16; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 52;
Matches   16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATCATGCGATATCGCA 16
          ||||||||||||||||
Db   159 ATCATGCGGATATCGCA 144

RESULT 5
LOCUS   AY102896
DEFINITION   Uncultured bacterium clone BCM3P-27B 16S ribosomal RNA gene,
             AY102896   745 bp   DNA   linear   BCT 24-JUN-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 12:49:40 ; Search time 1408.75 Seconds
(without alignments)
517.334 Million cell updates/sec

Title: SEQ4-4
Perfect score: 20
Sequence: 1 atcatgcgatatcatgcgat 20
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 32822875 seqs, 1821986908 residues
Word size : 0

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	80.0	569	7	CO010324	CO010324 EST798659
2	16	80.0	731	7	CO015681	CO015681 EST786063
3	16	80.0	770	7	CF812011	CF812011 EST689393
4	16	80.0	773	7	CO017170	CO017170 EST787552
5	16	80.0	798	7	CO034183	CO034183 EST812567
6	16	80.0	811	7	CO019122	CO019122 EST815215
7	16	80.0	864	7	CO029163	CO029163 EST807547
8	16	80.0	879	7	CO033025	CO033025 EST811409
9	16	80.0	927	7	CO020014	CO020014 EST816107
10	16	80.0	941	7	CO021731	CO021731 EST817824
11	16	80.0	944	7	CO032174	CO032174 EST810558
12	16	80.0	974	7	CO022893	CO022893 EST818986
13	16	80.0	1008	7	CO026836	CO026836 EST805220
14	16	80.0	1009	7	CO033329	CO033329 EST811713
15	16	80.0	156	7	CK099767	CK099767 A083P74.5
16	16	75.0	162	9	CG798199	CG798199 ZMMBBb034
17	16	75.0	264	1	AV292574	AV292574
18	16	75.0	271	5	BX765396	BX765396
19	16	75.0	278	4	BM582521	BM582521 17OOO6593
20	16	75.0	283	5	BX604726	BX604726
21	16	75.0	304	5	BX766921	BX766921
22	16	75.0	305	5	BX603239	BX603239
23	16	75.0	328	5	BX765460	BX765460
24	15	75.0	383	9	CC467667	CC467667 CH240_137
25	15	75.0	467	5	BX614015	BX614015 BX614015
26	15	75.0	517	5	BX606907	BX606907 BX606907
27	15	75.0	519	4	BM611622	BM611622 170006871
28	15	75.0	529	8	BH450732	BH450732 BOHCT76TR
29	15	75.0	531	8	BX766735	BX766735 BX766735
30	15	75.0	540	8	AZ145886	AZ145886 SP_0015_B
31	15	75.0	545	5	BX616446	BX616446 BX616446
32	15	75.0	546	8	BX614016	BX614016 BX614016
33	15	75.0	552	8	AZ306453	AZ306453 IM0007J01
34	15	75.0	555	4	BM590645	BM590645 170006873
35	15	75.0	556	5	BX604558	BX604558 BX604558
36	15	75.0	563	8	BZ415598	BZ415598 i557b05-g
37	15	75.0	568	8	AQ177776	AQ177776 HS_5507_B
38	15	75.0	588	8	BU717886	BU717886 SJM2BSD01
39	15	75.0	591	4	BM589323	BM589323 170006873
40	15	75.0	592	9	CG413516	CG413516 ZMMBBb027
41	15	75.0	602	9	CG799749	CG799749 111800t4D1
42	15	75.0	609	4	BM581510	BM581510 170006872
43	15	75.0	610	6	CD744877	CD744877 IRB7_D01
44	15	75.0	614	5	BX604307	BX604307 BX604307
45	15	75.0	622	4	BM657349	BM657349 170006874

ALIGNMENTS

RESULT 1
CO010324
LOCUS   CO010324   569 bp   mRNA   linear   EST 09-JUN-2004
DEFINITION EST798659 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
            kb Coccidioides posadasii cDNA clone CIEBG91 3' end, mRNA sequence.
ACCESSION   CO010324
VERSION   CO010324.1 GI:48517213
KEYWORDS   EST.
SOURCE   Coccidioides posadasii
ORGANISM   Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1 (bases 1 to 569)
AUTHORS   Gardner,M.J. and Cole,G.T.
TITLE   Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
JOURNAL   Unpublished (2003)
COMMENT   Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org.
FEATURES   Location/Qualifiers
     source   1..569
                     /organism="Coccidioides posadasii"
                     /mol_type="mRNA"
                     /strain="C735"
                     /db_xref="taxon:199306"
                     /clone="CIEBG91"
                     /dev_stage="spherules"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /clone_lib="Coccidioides posadasii spherule cDNA library,
                     0.4 to 2.3 kb"
                     /note="vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
                     Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
                     kb"
ORIGIN

Query Match   80.0%; Score 16; DB 7; Length 569;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATCATGCGATATCATG 16
           ||||||||||||||||
Db   38 ATCATGCGATATCATG 53

RESULT 2

```
LOCUS       CO015681                 731 bp    mRNA    linear   EST 09-JUN-2004
DEFINITION  CO015681 EST786063 Coccidioides posadasii saprobic phase cDNA library,
            greater than 4kb Coccidioides posadasii cDNA clone CIBBJ81 3' end,
            mRNA sequence.
ACCESSION   CO015681
VERSION     CO015681.1  GI:48522570
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
ORGANISM    Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 731)
AUTHORS     Gardner,M.J. and Cole,G.T.
TITLE       Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: EST786064
            Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org.
FEATURES             Location/Qualifiers
     source          1..731
                     /organism="Coccidioides posadasii"
                     /mol_type="mRNA"
                     /strain="C735"
                     /db_xref="taxon:199306"
                     /clone="CIBBJ81"
                     /dev_stage="saprobic phase (mycelia)"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /clone_lib="Coccidioides posadasii saprobic phase cDNA
                     library, greater than 4kb"
                     /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
                     Coccidioides posadasii saprobic phase cDNA library, size
                     fractionated cDNA > 4 kb"
ORIGIN

Query Match       80.0%;  Score 16;  DB 7;  Length 731;
Best Local Similarity  100.0%;  Pred. No. 8.7;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATCATGGGATATCATG 16
        ||||||||||||||||
Db  310 ATCATGGGATATCATG 325
```

RESULT 3

```
LOCUS       CF812011                 770 bp    mRNA    linear   EST 01-APR-2004
DEFINITION  CF812011 EST689393 Coccidioides posadasii saprobic phase cDNA library,
            greater than 4kb Coccidioides posadasii cDNA clone CIBA171 3' end,
            mRNA sequence.
ACCESSION   CF812011
VERSION     CF812011.1  GI:45917889
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
ORGANISM    Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 770)
AUTHORS     Gardner,M.J. and Cole,G.T.
TITLE       Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: EST689394
            Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org

Query Match       80.0%;  Score 16;  DB 7;  Length 770;
Best Local Similarity  100.0%;  Pred. No. 8.6;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATCATGGGATATCATG 16
        ||||||||||||||||
Db  176 ATCATGGGATATCATG 191
```

RESULT 4

```
LOCUS       CO017170                 773 bp    mRNA    linear   EST 09-JUN-2004
DEFINITION  CO017170 EST787552 Coccidioides posadasii saprobic phase cDNA library,
            greater than 4kb Coccidioides posadasii cDNA clone CIBBS86 5' end,
            mRNA sequence.
ACCESSION   CO017170
VERSION     CO017170.1  GI:48524059
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
ORGANISM    Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 773)
AUTHORS     Gardner,M.J. and Cole,G.T.
TITLE       Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: EST787551
            Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org
FEATURES             Location/Qualifiers
     source          1..773
                     /organism="Coccidioides posadasii"
                     /mol_type="mRNA"
                     /strain="C735"
                     /db_xref="taxon:199306"
                     /clone="CIBBS86"
                     /dev_stage="saprobic phase (mycelia)"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /clone_lib="Coccidioides posadasii saprobic phase cDNA
                     library, greater than 4kb"
                     /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
                     Coccidioides posadasii saprobic phase cDNA library, size
                     fractionated cDNA > 4 kb"
                     Seq primer: M13 Reverse.
ORIGIN

Query Match       80.0%;  Score 16;  DB 7;  Length 773;
Best Local Similarity  100.0%;  Pred. No. 8.6;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATCATGGGATATCATG 16
        ||||||||||||||||
Db  176 ATCATGGGATATCATG 191
```

```
Post-processing: Listing first 45 summaries

Database :   N_Geneseq_23Sep04:*
             1: geneseqn1980s:*
             2: geneseqn1990s:*
             3: geneseqn2000s:*
             4: geneseqn2001as:*
             5: geneseqn2001bs:*
             6: geneseqn2002as:*
             7: geneseqn2002bs:*
             8: geneseqn2003as:*
             9: geneseqn2003bs:*
            10: geneseqn2003cs:*
            11: geneseqn2003ds:*
            12: geneseqn2004s:*

             Pred. No. is the number of results predicted by chance to have a
             score greater than or equal to the score of the result being printed,
             and is derived by analysis of the total score distribution.
```

```
OM nucleic - nucleic search, using sw model
                                             Copyright (c) 1993 - 2004 Compugen Ltd.
                                             GenCore version 5.1.6

Run on:          October 27, 2004, 11:48:05 ; Search time 168.5 seconds
                                             (without alignments)
                                             623.077 Million cell updates/sec

Title:           SEQ4-4
Perfect score:   20
Sequence:        1 atcatgcgatatcatgcgat 20
Scoring table:   OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0
Searched:        4134886 seqs, 2624710521 residues
Word size :      0

Total number of hits satisfying chosen parameters:            8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

```
                                    SUMMARIES

Result                %     Query
 No.   Score   Match  Length   DB   ID        Description
-----  -----  -----  -------  ---  --------   -----------
  1      20   100.0    467     6   ABL39818   Abl39818 Human NS
  2      20   100.0    468     6   ABL39817   Abl39817 Human NS
c 3      14    70.0    162    12   ACH93501   Ach93501 Human gen
  4      14    70.0    417     8   ABX43931   Abx43931 Bovine ES
  5      14    70.0    510    12   ACH79801   Ach79801 Human gen
  6      14    70.0   1000     4   AAF91415   Aaf91415 Moraxella
  7      14    70.0   1000     6   ABK37795   Abk37795 DNA seque
  8      14    70.0   1029    10   ADB79861   Adb79861 Mouse SCC
c 9      14    70.0   1044    12   ADL02524   Adl02524 DNA encod
c 10     14    70.0   1086    12   ADL02896   Adl02896 DNA encod
  11     14    70.0   1578     4   AAH57572   Aah57572 Human bra
  12     14    70.0   1654    10   ABT42291   Abt42291 Toxicity
  13     14    70.0   1810    10   ADB62769   Adb62769 Human cDN
  14     14    70.0   1865     6   ABL39816   Abl39816 Human NS
  15     14    70.0   1902    10   ADF90621   Adf90621 Human hep
  16     14    70.0   2000    12   ADJ41057   Adj41057 Plant cDN
  17     14    70.0   2355    12   ADJ87643   Adj87643 Nervous s
c 18     14    70.0   2903    10   ADB56517   Adb56517 Rat gene
c 19     14    70.0   9095    10   ADB81174   Adb81174 MlcA codi
c 20     14    70.0   9099     6   AAS16809   Aas16809 Penicilli
c 21     14    70.0  14483     8   ACC48945   Acc48945 Escherich
  22     14    70.0  34203     4   AAF74546   Aaf74546 Penicilli
  23     14    70.0  34203     6   AAS16770   Aas16770 Penicilli
  24     14    70.0  34463     4   AAF74547   Aaf74547 Penicilli
  25     14    70.0  72149    10   ADB81173   Adb81173 MlcB s
  26     14    70.0  89047     4   AAF28547   Aaf28547 Genomic f
  27     14    70.0  99629     4   AAF28550   Aaf28550 Genomic f
  28     14    70.0 110000    10   ACF67367_36 Continuation (37 o
  29     14    70.0 110000    10   ACF65388_11 Continuation (12 o
c 30     14    70.0  27       10   AAV21928   Aav21928 Nuclease
c 31     14    65.0  239      12   ADL60151   Adl60151 Cat visua
c 32     13    65.0  270      12   ADI60195   Adi60195 Rat cDNA
  33     13    65.0  492       6   ABN92710   Abn92710 Staphyloc
  34     13    65.0  504       3   AAZ51989   Aaz51989 Salmonell
  35     13    65.0  574      12   ACH80026   Ach80026 Human gen
c 36     13    65.0  686       2   AAQ73205   Aaq73205 Gene codi
c 37     13    65.0  737       2   ADR02240   Adr02240 A. gossyp
  38     13    65.0  795      10   ABX07363   Abx07363 S. pneumo
  39     13    65.0  891       8   AAD35513   Aad35513 Pantoea s
  40     13    65.0  891       8   ACC44763   Acc44763 Pantoea s
  41     13    65.0  891      12   ADG14632   Adg14632 Pantoea s
  42     13    65.0  891      12   ADQ77207   Adq77207 Nucleotid
c 43     13    65.0  891      12   ADQ48642   Adq48642 Pantoea s
c 44     13    65.0  893       8   ABT14194   Abt14194 Pantoea s
c 45     13    65.0 1028       4   AAF68001   Aaf68001 Corynebac
```

```
                                    ALIGNMENTS

RESULT 1
ABL39818
ID    ABL39818 standard; cDNA; 467 BP.
XX
AC    ABL39818;
XX
DT    10-MAY-2002  (first entry)
XX
DE    Human NS cDNA sequence SEQ ID NO:128.
XX
KW    Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW    antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW    vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW    anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW    antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW    anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
KW    gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW    contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW    endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW    rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW    inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW    infertility; cardiovascular disease; coagulation disease; hypertension;
KW    ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW    diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW    gastric ulcer; Alzheimer's disease; gene; ss.
OS    Homo sapiens.
XX
PN    WO2020206315-A2.
XX
PD    24-JAN-2002.
XX
PF    17-JUL-2001; 2001WO-IL000653.
XX
PR    18-JUL-2000; 2000IL-00137345.
PR    15-DEC-2000; 2000IL-00140354.
XX
PA    (COMP-) COMPUGEN LTD.
XX
PI    Mintz L, Freilich S, Bernstein J;
XX
PT
XX
DR    WPI; 2002-155037/20.
DR    P-PSDB; ABB06164.
XX
```

Left column

PT   One hundred and twenty eight novel nucleic acid sequences, useful for
PT   treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
PS   Claim 1; Page 148; 290pp; English.
XX
CC   ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC   encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC   (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC   antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC   vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC   anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC   anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC   immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC   antidepressant, gastrointestinal, aneuroleptic, cerebroprotective,
CC   nootropic and contraceptive activities. The NS can be used in vaccines,
CC   gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC   antibodies from the present invention can be used for treating and
CC   diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC   diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC   cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC   glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC   disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC   disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC   depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC   Alzheimer's disease and as a contraceptive
XX
SQ   Sequence 467 BP; 144 A; 101 C; 112 G; 110 T; 0 U; 0 Other;

Query Match   100.0%; Score 20; DB 6; Length 467;
Best Local Similarity   100.0%;
          Matches   20; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATCATGGCGGATATCATGCGAT 20
              ||||||||||||||||||||||
Db   401 ATCATGGCGGATATCATGCGAT 420

RESULT 2
ABL39817
ID   ABL39817 standard; cDNA; 468 BP.
XX
AC   ABL39817;
XX
DT   10-MAY-2002 (first entry)
XX
DE   Human NS cDNA sequence SEQ ID NO:127.
XX
KW   Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW   antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW   vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW   anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW   antifibrinolytic; hypotension; antiasthmatic; cardiovascular;
KW   antiinvulsant; antidiabetic; immunomodulator; cardiant;
KW   anticonvulsant; virucide; antiulcer; cerebroprotective; nootropic;
KW   gastrointestinal; tranquilliser; antidepressant; aneuroleptic;
KW   contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW   endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW   rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW   inflammation; skin disorder; muscular dystrophy; hypertension;
KW   infertility; cardiovascular disease; coagulation disease; hypertension;
KW   ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW   diabetes; anxiety; depression; schizophrenia; angina; neurodegeneration;
KW   gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200206315-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-IL000653.
XX
PR   18-JUL-2000; 2000IL-00137345.
PR   15-DEC-2000; 2000IL-00140354.

Right column

XX
PA   (COMP-) COMPUGEN LTD.
XX
PI   Mintz L, Freilich S, Bernstein U;
XX
DR   WPI; 2002-155037/20.
DR   P-PSDB; ABB06163.
XX
PT   One hundred and twenty eight novel nucleic acid sequences, useful for
PT   treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
PS   Claim 1; Page 148; 290pp; English.
XX
CC   ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC   encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC   (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC   antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC   vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC   anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC   anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC   immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC   antidepressant, gastrointestinal, aneuroleptic, cerebroprotective,
CC   nootropic and contraceptive activities. The NS can be used in vaccines,
CC   gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC   antibodies from the present invention can be used for treating and
CC   diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC   diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC   cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC   glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC   disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC   disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC   depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC   Alzheimer's disease and as a contraceptive
XX
SQ   Sequence 468 BP; 144 A; 101 C; 113 G; 110 T; 0 U; 0 Other;

Query Match   100.0%; Score 20; DB 6; Length 468;
Best Local Similarity   100.0%;
          Matches   20; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATCATGGCGGATATCATGCGAT 20
              ||||||||||||||||||||||
Db   402 ATCATGGCGGATATCATGCGAT 421

RESULT 3
ACH93501
ID   ACH93501 standard; DNA; 162 BP.
XX
AC   ACH93501;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Human genome derived single exon probe #26696.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US20031947O4-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00028386.
XX
PR   03-APR-2002; 2002US-00028386.
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG, Rank DR, Hanzel DK;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 11:48:05 ; Search time 168.5 Seconds
           (without alignments)
           623.077 Million cell updates/sec

Title: SEQ1-C1
Perfect score: 20
Sequence: 1 atctcaggatatcctggagat 20
Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0
Searched: 4134886 seqs, 2624710521 residues
Word size :   0

Total number of hits satisfying chosen parameters: 8269772

Minimum D3 seq length: 0
Maximum D3 seq length: 200000000

Post-processing: Listing first 45 summaries

Database :   N_Geneseq_23Sep04:*
                1: geneseqn1980s:*
                2: geneseqn1990s:*
                3: geneseqn2000s:*
                4: geneseqn2001as:*
                5: geneseqn2001bs:*
                6: geneseqn2002as:*
                7: geneseqn2002bs:*
                8: geneseqn2003as:*
                9: geneseqn2003bs:*
               10: geneseqn2003cs:*
               11: geneseqn2003ds:*
               12: geneseqn2008s:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	17	85.0	3398	5	ABA20106	Aba20106 Human ner
2	17	85.0	3398	5	ABA20106	Aba20106 Human ner
3	17	85.0	3398	5	ABA20107	Aba20107 Human ner
4	17	85.0	3398	5	ABA20107	Aba20107 Human ner
5	15	75.0	1341	5	AAS65451	Aas65451 DNA encod
6	15	75.0	1341	5	AAS65451	Aas65451 DNA encod
7	15	75.0	3177	10	ADC60712	Adc60712 Human mac
8	15	75.0	3177	10	ADC60712	Adc60712 Human mac
9	15	75.0	3329	10	ADB62284	Adb62284 Human CDN
10	15	75.0	3329	10	ADB62284	Adb62284 Human CDN
11	15	75.0	5880	4	AAI68687	Aai68687 Pseudomon
12	15	75.0	5880	4	AAI68687	Aai68687 Pseudomon
13	15	75.0	6722	10	ADF81629	Adf81629 Leukaemia
14	15	75.0	6722	12	ADF81629	Adf81629 Leukaemia
15	15	75.0	6722	12	ADQ20617	Adq20617 Human sof
16	15	75.0	6722	12	ADQ20617	Adq20617 Human sof
17	15	75.0	8640	12	ADQ24600	Adq24600 Human sof
18	15	75.0	8640	12	ADQ24600	Adq24600 Human sof
19	15	75.0	11099	10	ADE25673	Ade25673 Human CDN
20	15	75.0	11099	10	ADE25673	Ade25673 Human CDN
21	15	75.0	11355	4	AAI68692	Aai68692 Pseudomon
22	15	75.0	11355	4	AAI68692	Aai68692 Pseudomon
23	15	75.0	57248	6	ABK83563	Abk83563 Human CDN
24	15	75.0	57248	6	ABK83563	Abk83563 Human CDN
25	14	70.0	465	12	ADL03996	Adl03996 DNA encod
26	14	70.0	465	12	ADL03996	Adl03996 DNA encod
27	14	70.0	640	3	AAZ80552	Aaz80552 Human col
28	14	70.0	640	3	AAZ80552	Aaz80552 Human col
29	14	70.0	696	9	ADA30915	Ada30915 DNA encod
30	14	70.0	696	9	ADA30915	Ada30915 DNA encod
31	14	70.0	1293	11	ACH97940	Ach97940 Klebsiell
32	14	70.0	1293	11	ACH97940	Ach97940 Klebsiell
33	14	70.0	1416	10	ACF67440	Acf67440 Photothab
34	14	70.0	1416	10	ACF67440	Acf67440 Photothab
35	14	70.0	1564	6	ABS55575	Abs55575 Human CDN
36	14	70.0	1564	6	ABS55575	Abs55575 Human CDN
37	14	70.0	1680	6	ABK86904	Abk86904 Human Ras
38	14	70.0	1680	6	ABK86904	Abk86904 Human Ras
39	14	70.0	1680	9	ADA02925	Ada02925 Human RAS
40	14	70.0	1680	9	ADA02925	Ada02925 Human RAS
41	14	70.0	1680	10	ADB72663	Adb72663 Human RAS
42	14	70.0	1680	10	ADB72663	Adb72663 Human RAS
43	14	70.0	1680	10	ADC85405	Adc85405 Human Ras
44	14	70.0	1680	10	ADC85405	Adc85405 Human Ras
45	14	70.0	1680	12	ADM74520	Adm74520 Human car

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
ABA20106
ID   ABA20106 standard; DNA; 3398 BP.
XX
AC   ABA20106;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 12437.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder; ds.
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine;
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0188974P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216680P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.

```
PR   14-AUG-2000;  2000US-0244518P.
PR   14-AUG-2000;  2000US-0244519P.
PR   08-NOV-2000;  2000US-0252513P.
PR   14-AUG-2000;  2000US-0255214P.
PR   14-AUG-2000;  2000US-0255266P.
PR   14-AUG-2000;  2000US-0255267P.
PR   14-AUG-2000;  2000US-0255268P.
PR   14-AUG-2000;  2000US-0255270P.
PR   14-AUG-2000;  2000US-0255447P.
PR   14-AUG-2000;  2000US-0255457P.
PR   14-AUG-2000;  2000US-0255758P.
PR   14-AUG-2000;  2000US-0255759P.
PR   18-AUG-2000;  2000US-0262679P.
PR   22-AUG-2000;  2000US-0226681P.
PR   22-AUG-2000;  2000US-0226868P.
PR   22-AUG-2000;  2000US-0227182P.
PR   23-AUG-2000;  2000US-0227009P.
PR   30-AUG-2000;  2000US-0288924P.
PR   01-SEP-2000;  2000US-0229287P.
PR   01-SEP-2000;  2000US-0229343P.
PR   01-SEP-2000;  2000US-0229344P.
PR   01-SEP-2000;  2000US-0229345P.
PR   05-SEP-2000;  2000US-0229509P.
PR   05-SEP-2000;  2000US-0229513P.
PR   06-SEP-2000;  2000US-0230437P.
PR   06-SEP-2000;  2000US-0230438P.
PR   08-SEP-2000;  2000US-0231242P.
PR   08-SEP-2000;  2000US-0231243P.
PR   08-SEP-2000;  2000US-0231244P.
PR   08-SEP-2000;  2000US-0231413P.
PR   08-SEP-2000;  2000US-0231414P.
PR   14-SEP-2000;  2000US-0232080P.
PR   08-SEP-2000;  2000US-0232081P.
PR   12-SEP-2000;  2000US-0231968P.
PR   14-SEP-2000;  2000US-0232397P.
PR   14-SEP-2000;  2000US-0232398P.
PR   14-SEP-2000;  2000US-0232399P.
PR   14-SEP-2000;  2000US-0232400P.
PR   14-SEP-2000;  2000US-0232401P.
PR   14-SEP-2000;  2000US-0233063P.
PR   14-SEP-2000;  2000US-0233064P.
PR   14-SEP-2000;  2000US-0233065P.
PR   21-SEP-2000;  2000US-0234223P.
PR   21-SEP-2000;  2000US-0234274P.
PR   25-SEP-2000;  2000US-0234997P.
PR   25-SEP-2000;  2000US-0234998P.
PR   26-SEP-2000;  2000US-0235834P.
PR   27-SEP-2000;  2000US-0235836P.
PR   27-SEP-2000;  2000US-0235834P.
PR   29-SEP-2000;  2000US-0236327P.
PR   29-SEP-2000;  2000US-0236367P.
PR   29-SEP-2000;  2000US-0236368P.
PR   29-SEP-2000;  2000US-0236369P.
PR   29-SEP-2000;  2000US-0236370P.
PR   02-OCT-2000;  2000US-0236802P.
PR   02-OCT-2000;  2000US-0237037P.
PR   02-OCT-2000;  2000US-0237038P.
PR   02-OCT-2000;  2000US-0237039P.
PR   02-OCT-2000;  2000US-0237040P.
PR   13-OCT-2000;  2000US-0239935P.
PR   13-OCT-2000;  2000US-0239937P.
PR   20-OCT-2000;  2000US-0240960P.
PR   20-OCT-2000;  2000US-0241785P.
PR   20-OCT-2000;  2000US-0241786P.
PR   20-OCT-2000;  2000US-0241787P.
PR   20-OCT-2000;  2000US-0241808P.
PR   20-OCT-2000;  2000US-0241809P.
PR   20-OCT-2000;  2000US-0242221P.
PR   20-OCT-2000;  2000US-0242222P.
PR   01-NOV-2000;  2000US-0244617P.
PR   08-NOV-2000;  2000US-0244674P.
PR   08-NOV-2000;  2000US-0244752P.
PR   08-NOV-2000;  2000US-0246476P.

PR   08-NOV-2000;  2000US-0246477P.
PR   08-NOV-2000;  2000US-0246478P.
PR   08-NOV-2000;  2000US-0246523P.
PR   08-NOV-2000;  2000US-0246524P.
PR   08-NOV-2000;  2000US-0246525P.
PR   08-NOV-2000;  2000US-0246526P.
PR   08-NOV-2000;  2000US-0246527P.
PR   08-NOV-2000;  2000US-0246528P.
PR   08-NOV-2000;  2000US-0246532P.
PR   08-NOV-2000;  2000US-0246609P.
PR   08-NOV-2000;  2000US-0246610P.
PR   08-NOV-2000;  2000US-0246611P.
PR   08-NOV-2000;  2000US-0246613P.
PR   17-NOV-2000;  2000US-0492207P.
PR   17-NOV-2000;  2000US-0492208P.
PR   17-NOV-2000;  2000US-0492209P.
PR   17-NOV-2000;  2000US-0492210P.
PR   17-NOV-2000;  2000US-0492211P.
PR   17-NOV-2000;  2000US-0492212P.
PR   17-NOV-2000;  2000US-0492213P.
PR   17-NOV-2000;  2000US-0492214P.
PR   17-NOV-2000;  2000US-0492215P.
PR   17-NOV-2000;  2000US-0492216P.
PR   17-NOV-2000;  2000US-0492172P.
PR   17-NOV-2000;  2000US-0492244P.
PR   17-NOV-2000;  2000US-0492445P.
PR   17-NOV-2000;  2000US-0492644P.
PR   17-NOV-2000;  2000US-0492652P.
PR   17-NOV-2000;  2000US-0492977P.
PR   17-NOV-2000;  2000US-0492999P.
PR   17-NOV-2000;  2000US-0493002P.
PR   01-DEC-2000;  2000US-0250391P.
PR   01-DEC-2000;  2000US-0251160P.
PR   05-DEC-2000;  2000US-0251030P.
PR   05-DEC-2000;  2000US-0251988P.
PR   05-DEC-2000;  2000US-0256719P.
PR   06-DEC-2000;  2000US-0251479P.
PR   08-DEC-2000;  2000US-0251856P.
PR   08-DEC-2000;  2000US-0251868P.
PR   08-DEC-2000;  2000US-0251869P.
PR   08-DEC-2000;  2000US-0251989P.
PR   08-DEC-2000;  2000US-0251990P.
PR   11-DEC-2000;  2000US-0254097P.
PR   05-JAN-2001;  2001US-0259678P.
```

XX

PS   Disclosure; SEQ ID NO 12437; 1701pp + Sequence Listing; English.

XX

PA   (HUMA-) HUMAN GENOME SCI INC.

XX

PI   Rosen CA, Barash SC, Ruben SM;

XX

DR   WPI; 2001-541565/60.

XX

PT   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating nervous system cancers
PT   and metastases.

XX

CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: October 27, 2004, 11:49:55 ; Search time 710.125 Seconds
(without alignments)
1331.870 Million cell updates/sec

```
Title:          SEQ1-C1
Perfect score:  20
Sequence:       1 atctcaggatatcctgagat 20
Scoring table:
                OLIGO_NUC
                Gapop 60.0 , Gapext 60.0
Searched:       4526729 seqs, 23644849745 residues
Word size :     0

Total number of hits satisfying chosen parameters:    9053458
Minimum DB seq length:      0
Maximum DB seq length:      200000000

Post-processing: Listing first 45 summaries

Database :
                GenEmbl:*
          1: gb_ba:*
          2: gb_htg:*
          3: gb_in:*
          4: gb_om:*
          5: gb_ov:*
          6: gb_pat:*
          7: gb_ph:*
          8: gb_pl:*
          9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB	ID	Description
c 1	19	95.0	251124	9	HUAE0000660	HUAE0000660 Homo sapi
c 2	19	95.0	251124	9	HUAE0000660	AE000660 Homo sapi
c 3	18	90.0	171263	10	AC118702	AC118702 Mus muscu
c 4	18	90.0	171263	10	AC118702	AC118702 Mus muscu
c 5	18	90.0	234203	2	AC121311	AC121311 Mus muscu
c 6	18	90.0	234203	2	AC121311	AC121311 Mus muscu
c 7	17	85.0	6172	1	MTMCR	X07794 M. thermoau
c 8	17	85.0	6172	1	MTMCR	X07794 M. thermoau
c 9	17	85.0	97011	9	HS997O15	AL035073 Human DNA
c 10	17	85.0	97011	9	HS997O15	AL035073 Human DNA
c 11	17	85.0	134995	2	AC122494	AC122494 Mus muscu
c 12	17	85.0	134995	2	AC122494	AC122494 Mus muscu
c 13	17	85.0	136839	9	AL139397	AL139397 Human DNA
c 14	17	85.0	136839	9	AL139397	AL139397 Human DNA
c 15	17	85.0	161266	10	AC127298	AC127298 Mus muscu
c 16	17	85.0	161266	10	AC127298	AC127298 Mus muscu
c 17	17	85.0	173849	9	AC078855	AC078855 Homo sapi
c 18	17	85.0	173849	9	AC078855	AC078855 Homo sapi
c 19	17	85.0	175464	9	AC069021	AC069021 Homo sapi
c 20	17	85.0	175464	9	AC069021	AC069021 Homo sapi
c 21	17	85.0	179481	2	AC148908	AC148908 Otolemur
c 22	17	85.0	179481	2	AC148908	AC148908 Otolemur
c 23	17	85.0	198756	9	AC087575	AC087575 Homo sapi
c 24	17	85.0	198756	9	AC087575	AC087575 Homo sapi
c 25	17	85.0	199035	2	AC022263	AC022263 Homo sapi
c 26	17	85.0	199035	2	AC022263	AC022263 Homo sapi
c 27	17	85.0	201708	2	AC020731	AC020731 Homo sapi
c 28	17	85.0	201708	2	AC020731	AC020731 Homo sapi
c 29	17	85.0	216523	9	AC073587	AC073587 Homo sapi
c 30	17	85.0	216523	9	AC073587	AC073587 Homo sapi
c 31	17	85.0	243707	2	AC126141	AC126141 Rattus no
c 32	17	85.0	243707	2	AC126141	AC126141 Rattus no
c 33	16	80.0	24808	9	HSN12G10	Z69712 Human DNA s
c 34	16	80.0	24808	9	HSN12G10	Z69712 Human DNA s
c 35	16	80.0	49105	10	AL731716	AL731716 Mouse DNA
c 36	16	80.0	49105	10	AL731716	AL731716 Mouse DNA
c 37	16	80.0	89626	9	AC004776	AC004776 Homo sapi
c 38	16	80.0	89626	9	AC004776	AC004776 Homo sapi
c 39	16	80.0	122356	10	AL845310	AL845310 Mouse DNA
c 40	16	80.0	122356	10	AL845310	AL845310 Mouse DNA
c 41	16	80.0	128980	2	AL139107	AL139107 Homo sapi
c 42	16	80.0	128980	2	AL139107	AL139107 Homo sapi
c 43	16	80.0	137493	10	AL732581	AL732581 Mouse DNA
c 44	16	80.0	137493	10	AL732581	AL732581 Mouse DNA
c 45	16	80.0	141324	10	AL645466	AL645466 Mouse DNA

ALIGNMENTS

```
RESULT 1
HUAE0000660
LOCUS       HUAE0000660      251124 bp    DNA     linear   PRI 19-DEC-2002
DEFINITION  Homo sapiens T-cell receptor alpha delta locus from bases 501613 to
            752736 (section 3 of 5) of the Complete Nucleotide Sequence.
ACCESSION   AE000660 AE000521 U85197
VERSION     AE000660.1  GI:2358042
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 251124)
AUTHORS     Koop,B.F., Rowen,L., Wang,K., Kuo,C.L., Seto,D., Lenstra,J.A.,
            Howard,S., Shan,W., Deshpande,P. and Hood,L.
TITLE       The human T-cell receptor TCRAC/TCRDC (C alpha/C delta) region:
            organization, sequence, and evolution of 97.6 Kb of DNA
JOURNAL     Genomics 19 (3), 478-493 (1994)
MEDLINE     94245236
PUBMED      8188290
REMARK      This citation covers from bases 966383-1064019
REFERENCE   2  (bases 1 to 251124)
AUTHORS     Boysen,C., Simon,M.I. and Hood,L.
TITLE       Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with
            bacterial artificial chromosome clones
JOURNAL     Genome Res. 7 (4), 330-338 (1997)
MEDLINE     97264339
PUBMED      9110172
REFERENCE   3  (bases 1 to 251124)
AUTHORS     Boysen,C., Inyoul,L., Smith,T.M., Smith,A., Wang,K., Rowen,L. and
            Hood,L.
TITLE       T-Cell Receptor Alpha Delta Locus Complete Nucleotide Sequence
JOURNAL     Unpublished
REMARK      This citation covers bases 1-983545 and bases 1064020-1071650
REFERENCE   4  (bases 1 to 251124)
AUTHORS     Boysen,C., Lee,I., Smith,T.M., Smith,A., Kai,W., Lee,R. and Leroy,H.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUL-1997) Department of Molecular Biotechnology,
            University of Washington, Box 357730, Seattle, Washington 98195,
            USA
REMARK      Complete nucleotide sequence of the human T-cell receptor alpha
            delta locus
```

```
COMMENT     On Jun 12, 2001 this sequence version replaced gi:2121231.
            Sequencing method: high redundancy shotgun.  Interspersed Repeats
            were identified with RepeatMasker (available from
            http://ftp.genome.washington.edu/RM/RepeatMasker.html) Simple
            sequence repeats were identified with sputnik (available from
            http://serac.mbt.washington.edu/~chrisa/software/sputnik.html).
FEATURES             Location/Qualifiers
     source          1..251124
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Mel Simon's BAC library"
                     /clone="BAC 480"
     source          157356..235794
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Mel Simon's BAC library"
                     /clone="BAC 378"
                     /note="This span of sequence was submitted independently
                     as Genbank Accession Number U85197."
     source          233572..251124
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Mel Simon's BAC library"
                     /clone="BAC 378"
                     /note="This span of sequence was submitted independently
                     as Genbank Accession Number U85196."
     repeat_region   complement(326..479)
                     /rpt_family="MIR"
     repeat_region   603..983
                     /rpt_family="MLT1B"
     gene            1738..2478
                     /gene="TCRAV17S1"
     promoter        1738..1757
                     /gene="TCRAV17S1"
     mRNA            join(1957..2008,2155..>2439)
                     /gene="TCRAV17S1"
                     /note="Old name TCRAV3S1"
     CDS             join(1957..2008,2155..>2439)
                     /gene="TCRAV17S1"
                     /note="Old name TCRAV3S1"
                     /codon_start=1
                     /product="TCRAV17S1"
                     /protein_id="AAB69018.1"
                     /db_xref="GI:2358043"
                     /translation="METLLGVSLVIWLQLARVNSQQGEEDPQALSIQEGENATWNCS
                     YKTSINNLQWYRQNSGRGLVHLIIRSNEREKHSGRLRVTLDTSKKSSSLLITASRAA
                     DTASYFCATD"
     v_segment       join(1957..2008,2155..>2439)
                     /gene="TCRAV17S1"
                     /standard_name="TCRAV17S1"
     misc_recomb     2440..2446
                     /gene="TCRAV17S1"
                     /note="heptamer"
     misc_recomb     2447..2469
                     /gene="TCRAV17S1"
                     /note="spacer"
     misc_recomb     2470..2478
                     /gene="TCRAV17S1"
                     /standard_name="TCRAV17S1"
                     /note="nonamer"
     repeat_region   complement(4306..4461)
                     /rpt_family="MIR2"
     repeat_region   4975..5275
                     /rpt_family="AluY"
     repeat_region   complement(5366..5574)
                     /rpt_family="MER58A"
     gene            7244..7983
                     /gene="TCRAV18S1"
                     /pseudo
                     /product="TCRAV18S1"
     mRNA            join(<7482..7527,7656..>7944)
                     /gene="TCRAV18S1"
                     /pseudo
                     /product="TCRAV18S1"
                     /standard_name="TCRAV18S1"
     promoter        7244..7263
                     /gene="TCRAV18S1"
                     /pseudo
                     /standard_name="TCRAV18S1"
     v_segment       join(7482..7527,7656..>7944)
                     /gene="TCRAV18S1"
                     /note="heptamer"
                     /pseudo
                     /standard_name="TCRAV18S1"
     misc_recomb     7945..7951
                     /gene="TCRAV18S1"
                     /standard_name="TCRAV18S1"
                     /note="spacer"
     misc_recomb     7952..7974
                     /gene="TCRAV18S1"
                     /standard_name="TCRAV18S1"
                     /note="spacer"
     misc_recomb     7975..7983
                     /gene="TCRAV18S1"
                     /standard_name="TCRAV18S1"
                     /note="nonamer"
     repeat_region   complement(8236..8389)
                     /rpt_family="MER69A"
     repeat_region   complement(8556..8740)
                     /rpt_family="L1MB7"
     repeat_region   complement(9281..10170)
                     /rpt_family="L1PA2"
     repeat_region   complement(10023..11230)
                     /rpt_family="L1L"
     misc_recomb     11234..11405
                     /gene="TCRAV19S1"
                     /note="nonamer"
     gene            11818..12491
                     /gene="TCRAV19S1"
     promoter        11818..11837
                     /gene="TCRAV19S1"
     mRNA            join(<11900..11948,12152..>12452)
                     /gene="TCRAV19S1"
                     /note="Old name TCRAV12S1"
     CDS             join(<11900..11948,12152..>12452)
                     /gene="TCRAV19S1"
                     /note="Old name TCRAV12S1"
                     /codon_start=1
                     /product="TCRAV19S1"
                     /protein_id="AAB69019.1"
                     /db_xref="GI:2358044"
                     /translation="MLTASLLRAVIASICVVSSMAQKVTQAQTEISVVEKEDVTLDCV
                     YETRDTYYLFWYKQPPSGELVFLIRRNSFDEQNEISGRYSWNFQKSTSSFNFTITAS
                     QVVDSAVYFCALSEA"
     v_segment       join(<11900..11948,12152..>12452)
                     /gene="TCRAV19S1"
                     /standard_name="TCRAV19S1"
     misc_recomb     12453..12459
                     /gene="TCRAV19S1"
                     /note="Old name TCRAV12S1"
     misc_recomb     12460..12482
                     /gene="TCRAV19S1"
                     /standard_name="TCRAV19S1"
                     /note="heptamer"
     misc_recomb     12483..12491
                     /gene="TCRAV19S1"
```

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 11:49:55 ; Search time 710.125 Seconds
                (without alignments)
                1331.870 Million cell updates/sec

Title: SEQ3-3
Perfect score: 20
Sequence: 1 atcggatgatatcgcatgat 20
Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0
Searched: 4526729 seqs, 2364849745 residues
Word size :   0
Total number of hits satisfying chosen parameters: 9053458
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   GenEmbl:*
                1: gb_ba:*
                2: gb_bt:*
                3: gb_in:*
                4: gb_om:*
                5: gb_ov:*
                6: gb_pat:*
                7: gb_ph:*
                8: gb_pl:*
                9: gb_pr:*
                10: gb_ro:*
                11: gb_sts:*
                12: gb_sy:*
                13: gb_un:*
                14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB ID	Description
1	20	100.0	467	6	AX364977	AX364977 Sequence
c 2	20	100.0	468	6	AX364976	AX364976 Sequence
c 3	20	100.0	239244	9	S82024	S82024 SCG10=neuro
4	16	80.0	696	9	AY580688	AY580688 Unculture
5	16	80.0	787	1	AP006500_01	Continuation (2 of
c 6	16	80.0	110000	2	AP006500_01	Continuation (2 of
c 7	16	80.0	116073	8	AP004027	AP004027 Oryza sat
c 8	16	80.0	187091	2	BX548161	BX548161 Danio rer
9	15	75.0	1405	1	AY082459	AY082459 Uncysture
c 10	15	75.0	2047	3	BMER82DNA	X58449 B.mori ErB.
c 11	15	75.0	60631	9	HS376H23	AL031078 Human DNA
12	15	75.0	145968	2	BX293554_4	Continuation (5 of
13	15	75.0	158127	9	AC141842	AC141842 Apis mell
14	15	75.0	166770	2	AC105394	AC105394 Homo sapi
15	15	75.0	18564	5	AC119966	AC119966 Mus muscu
c 16	15	75.0	18608	4	AL772314	AL772314 Zebrafish
c 17	15	75.0	18613	8	BX470181	BX470181 Zebrafish
c 18	15	75.0	18921	10	AP005689	AP005689 Oryza sat
c 19	15	75.0	193706	10	AL929443	AL929443 Mouse DNA
c 20	15	75.0	210310	5	BX005322	BX005322 Zebrafish
c 21	15	75.0	211084	9	AL050308	AL050308 Human DNA
c 22	15	75.0	232367	2	AC145242	AC145242 Gorilla 9
c 23	15	75.0	239244	10	AL606925	AL606925 Mouse DNA
24	15	75.0	259013	2	AC096892	AC096892 Rattus no
25	15	75.0	265401	2	AC099177	AC099177 Rattus no
26	15	75.0	275519	2	AC119596	AC119596 Rattus no
c 27	14	70.0	643	8	AF349708	AF349708 Russula s
c 28	14	70.0	656	5	X71433	X71433 X.laevis SC
c 29	14	70.0	724	9	BC006302	BC006302 Homo sapi
c 30	14	70.0	732	5	BC077672	BC077672 Xenopus t
c 31	14	70.0	772	9	D50375	D50375 Homo sapien
c 32	14	70.0	783	5	XLSCG10B	X71434 X.laevis SC
c 33	14	70.0	885	5	L14938	L14938 Chicken SCG
c 34	14	70.0	1000	6	AX081536	AX081536 Sequence
c 35	14	70.0	1000	6	AX374717	AX374717 Sequence
c 36	14	70.0	1029	6	AX683127	AX683127 Sequence
c 37	14	70.0	1029	10	L20263	L20263 Mouse SCG10
c 38	14	70.0	1044	6	AR449546	AR449546 Sequence
c 39	14	70.0	1086	1	AR449918	AR449918 Sequence
c 40	14	70.0	1479	1	AY532556	AY532556 Uncysture
c 41	14	70.0	1488	9	AL110174	AL110174 Homo sapi
c 42	14	70.0	1494	1	UGA318170	AJ318170 Uncysture
c 43	14	70.0	1507	1	BC056321	BC056321 Danio rer
c 44	14	70.0	1554	10	AF306458	AF306458 Rattus no
c 45	14	70.0	1799	5	BC075248	BC075248 Xenopus 1

ALIGNMENTS

RESULT 1
AX364977/c
LOCUS   AX364977   467 bp   DNA   linear   PAT 15-FEB-2002
DEFINITION Sequence 128 from Patent WO0206315.
ACCESSION   AX364977
VERSION   AX364977.1   GI:18696867
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mintz,L., Freilich,S. and Bernstein,J.
  TITLE   Novel nucleic acid and amino acid sequences
  JOURNAL   Patent: WO 0206315-A 128 24-JAN-2002;
            Compugen Ltd. (IL)
FEATURES   Location/Qualifiers
     source   1..467
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%; Score 20; DB 6; Length 467;
Best Local Similarity   100.0%;
Matches   20; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATCGGATGATATCGCATGAT 20
              ||||||||||||||||||||
Db   420 ATCGGATGATATCGCATGAT 401

RESULT 2
AX364976/c
LOCUS   AX364976   468 bp   DNA   linear   PAT 15-FEB-2002
DEFINITION Sequence 127 from Patent WO0206315.
ACCESSION   AX364976
VERSION   AX364976.1   GI:18696866
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```
REFERENCE   1
  AUTHORS   Mintz,L., Freilich,S. and Bernstein,J.
  TITLE     Novel nucleic acid and amino acid sequences
  JOURNAL   Patent: WO 0206315-A 127 24-JAN-2002;
            Compugen Ltd. (IL)
FEATURES             Location/Qualifiers
     source          1..468
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Qy        1 ATCGCATGATATCGCCATGAT 20
            |||||||||||||||||||||
Db      421 ATCGCATGATATCGCCATGAT 402

Query Match        100.0%;  Score 20;  DB 6;  Length 468;
Best Local Similarity  100.0%;  Pred. No. 0.19;
Matches   20;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 3
S82024/c
```
LOCUS       S82024                   696 bp    mRNA    linear   PRI 03-AUG-1996
DEFINITION  SCG10=neuron-specific growth-associated protein/stathmin homolog
            [human, embryo, mRNA, 696 nt].
ACCESSION   S82024
VERSION     S82024.1  GI:1478502
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 696)
  AUTHORS   Okazaki,T., Wang,H., Masliah,E., Cao,M., Johnson,S.A., Sundsmo,M.,
            Saitoh,T. and Mori,N.
  TITLE     SCG10, a neuron-specific growth-associated protein in Alzheimer's
            disease
  JOURNAL   Neurobiol. Aging 16 (6), 883-894 (1995)
  MEDLINE   96192979
  PUBMED    8622778
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 177683] from the original journal article.
FEATURES             Location/Qualifiers
     source          1..696
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1..696
                     /gene="SCG10"
     CDS             29..568
                     /gene="SCG10"
                     /codon_start=1
                     /product="SCG10"
                     /note="neuron-specific growth-associated protein/stathmin
                     homolog"
                     /protein_id="AAB36428.1"
                     /db_xref="GI:1478503"
                     /translation="MAKTAMAYKEKMKELSMLSLICSCFYPERNINIYTYDDMEVKQ
                     INKRASGQAFELILIKPSPISRAPRTLASPKKKDLSLEEIQKKLEAAGERKSQEAQV
                     LKQLAEKREHEREVLQVALEENNFSKVAEEKLILKMEQIKENREANLAAIIERLQEK
                     ERHAAEVRRNKELQVELSG"
ORIGIN

Qy        1 ATCGCATGATATCGCCATGAT 20
            |||||||||||||||||||||
Db      649 ATCGCATGATATCGCCATGAT 630

Query Match        100.0%;  Score 20;  DB 9;  Length 696;
Best Local Similarity  100.0%;  Pred. No. 0.18;
Matches   20;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 4
AY580688
```
LOCUS       AY580688                 787 bp    DNA     linear   BCT 29-JUL-2004
DEFINITION  Uncultured Bacteroidetes bacterium clone PI_4e1og 16S ribosomal RNA
            gene gene, partial sequence.
ACCESSION   AY580688
VERSION     AY580688.1  GI:46392307
KEYWORDS    .
SOURCE      uncultured Bacteroidetes bacterium
  ORGANISM  uncultured Bacteroidetes bacterium
            Bacteria; Bacteroidetes; environmental samples.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   Acinas,S.G., Klepac-Ceraj,V., Hunt,D.E., Pharino,C., Ceraj,I.,
            Distel,D.L. and Polz,M.F.
  TITLE     Fine-scale phylogenetic architecture of a complex bacterial
            community
  JOURNAL   Nature 430 (6999), 551-554 (2004)
  PUBMED    15282603
REFERENCE   2  (bases 1 to 787)
  AUTHORS   Acinas,S.G., Klepac-Ceraj,V., Hunt,D.E., Pharino,C., Ceraj,I.,
            Distel,D.L. and Polz,M.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2004) Department of Civil and Environmental
            Engineering, Massachusetts Institute of Technology (MIT), 15 Vassar
            St., Room 48-108, Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1..787
                     /organism="uncultured Bacteroidetes bacterium"
                     /mol_type="genomic DNA"
                     /isolation_source="coastal bacterioplankton sample of Plum
                     Island Sound Estuary"
                     /db_xref="taxon:152509"
                     /clone="PI_4e1og"
                     /environmental_sample
                     /country="USA: northeastern Massachusetts"
     rRNA            <1..>787
                     /product="16S ribosomal RNA gene"
ORIGIN

Qy        5 CATGATATCGCCATGAT 20
            |||||||||||||||||
Db      143 CATGATATCGCCATGAT 158

Query Match         80.0%;  Score 16;  DB 1;  Length 787;
Best Local Similarity  100.0%;  Pred. No. 42;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 5
AP006500
AP006500_01
WPCOMMENT
WPCOMMENT
```
Sequence split into 13 fragments      LOCUS AP006500 Accession AP006500
   Fragment Name       Begin          End
   AP006500_00             1       110000
   AP006500_01        100001       210000
   AP006500_02        200001       310000
   AP006500_03        300001       410000
   AP006500_04        400001       510000
   AP006500_05        500001       610000
   AP006500_06        600001       710000
   AP006500_07        700001       810000
   AP006500_08        800001       910000
   AP006500_09        900001      1010000
   AP006500_10       1000001      1110000
   AP006500_11       1100001      1210000
   AP006500_12       1200001      1254097
Continuation (2 of 13) of AP006500 from base 100001 (AP006500 Cyanidioschyzon merolae ;

Qy        1 ATCGCATGATATCGCCATGAT 20
            |||||||||||||||||||||
Db      ...  ATCGCATGATATCGCCATGAT

Query Match         80.0%;  Score 16;  DB 2;  Length 110000;
Best Local Similarity  100.0%;  Pred. No. 30;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 12:49:40 ; Search time 1408.75 Seconds
                                            (without alignments)
                                            517.334 Million cell updates/sec

Title: SEQ2-C2
Perfect score: 20
Sequence: 1 atcagtcgatatcgactgat 20
Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0
Searched: 32822875 seqs, 18219865908 residues
Word size :   0

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries
Database :

               EST:*
               1: gb_est1:*
               2: gb_est2:*
               3: gb_htc:*
               4: gb_est3:*
               5: gb_est4:*
               6: gb_est5:*
               7: gb_et6:*
               8: gb_gss1:*
               9: gb_gss2:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

                                   %
Result   Query
No.   Score Match Length DB ID   Description
------ ----- ----- ------ -- ---------- -----------

  1   18   90.0   797   9   CG137880   CG137880 PUJEX65TB
  2   18   90.0   797   9   CG137880   CG137880 PUJEX65TB
  3   18   90.0   946   9   CG069001   CG069001 PUFNB43TD
  4   18   90.0   946   9   CG069001   CG069001 PUFNB43TD
  5   18   90.0   1005   9   CG069000   CG069000 PUFNB43TB
  6   18   90.0   1005   9   CG069000   CG069000 PUFNB43TB
  7   15   75.0   346   9   DR38G4S   AL981527 Danio rer
  8   15   75.0   346   9   DR38G4S   AL981527 Danio rer
  9   15   75.0   354   8   CC098415   CC098415 CSU-K34.1
 10   15   75.0   354   8   CC098415   CC098415 CSU-K34.1
 11   15   75.0   473   8   AZ048828   AZ048828 GSSBru022
 12   15   75.0   473   8   AZ048828   AZ048828 GSSBru022
 13   15   75.0   603   8   CC067257   CC067257 CSU-K33r.
 14   15   75.0   603   8   CC067257   CC067257 CSU-K33r.
 15   15   75.0   651   7   CO358321   CO358321 DR_ATE_OR
 16   15   75.0   651   7   CO358321   CO358321 DR_ATE_OR
 17   15   75.0   762   9   BX221404   BX221404 Danio rer
 18   15   75.0   762   9   BX221404   BX221404 Danio rer
 19   15   75.0   770   8   CC075416   CC075416 CSU-K33r.
 20   15   75.0   770   8   CC075416   CC075416 CSU-K33r.
 21   15   75.0   785   9   BX169079   BX169079 Danio rer
 22   15   75.0   785   9   BX169079   BX169079 Danio rer
 23   15   75.0   787   9   BX213610   BX213610 Danio rer
 24   15   75.0   787   9   BX213610   BX213610 Danio rer
 25   15   75.0   801   7   CN209370   CN209370 Tor9883 G
 26   15   75.0   801   7   CN209370   CN209370 Tor9883 G
 27   15   75.0   849   9   BX190329   BX190329 Danio rer
 28   15   75.0   849   9   BX190329   BX190329 Danio rer
 29   14   70.0   181   9   CG101354   CG101354 PUFSC29TB
 30   14   70.0   181   9   CG101354   CG101354 PUFSC29TB
 31   14   70.0   261   7   D65272   D65272 CELKI33H4R
 32   14   70.0   261   7   D65272   D65272 CELKI33H4R
 33   14   70.0   305   8   B74261   B74261 CIT-HSP-203
 34   14   70.0   305   8   B74261   B74261 CIT-HSP-203
 35   14   70.0   329   9   CG101355   CG101355 PUFSC29TD
 36   14   70.0   329   9   CG101355   CG101355 PUFSC29TD
 37   14   70.0   358   5   BU640957   BU640957 mgnk003xD
 38   14   70.0   358   5   BU640957   BU640957 mgnk003xD
 39   14   70.0   378   2   BE612031   BE612031 sr02h10.Y
 40   14   70.0   378   2   BE612031   BE612031 sr02h10.Y
 41   14   70.0   413   8   BH900567   BH900567 Otb00109
 42   14   70.0   413   8   BH900567   BH900567 Otb00109
 43   14   70.0   427   8   AQ047442   AQ047442 CLM-1a9-u
 44   14   70.0   427   8   AQ047442   AQ047442 CLM-1a9-u
 45   14   70.0   446   6   CF120928   CF120928 TgsSTzyi2

ALIGNMENTS

RESULT 1
LOCUS   CG137880   797 bp   DNA   linear   GSS 21-AUG-2003
DEFINITION PUJEX65TB ZM_0.6_1.0 KB Zea mays genomic clone ZMMBTa0652L10,
            genomic survey sequence.
ACCESSION   CG137880
VERSION   CG137880.1 GI:34028663
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 797)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT   Other_GSSs: PUJEX65TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: IR
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..797
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTac0655L10"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CcF selected genomic DNA library"
ORIGIN

            Query Match   90.0%; Score 18; DB 9; Length 797;
            Best Local Similarity 100.0%; Pred. No. 0.051;
            Matches   18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATCAGTCGATATCGACTG 18
              ||||||||||||||||||
Db   349 ATCAGTCGATATCGACTG 366

RESULT 2
LOCUS   CG137880
DEFINITION   PUEX65TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0652L10,
                genomic survey sequence.
ACCESSION   CG137880
VERSION   CG137880.1 GI:34028663
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 797)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
                Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
                Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
  COMMENT   Other_GSSs: PUEX65TD
                Contact: Cathy Whitelaw
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA
                Tel: 301-838-5843
                Fax: 301-838-0208
                Email: whitelaw@tigr.org
                Seq primer: TR
                Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..797
                 /organism="zea mays"
                 /mol_type="genomic DNA"
                 /strain="B73"
                 /db_xref="taxon:4577"
                 /clone_lib="ZM_0.6_1.0_KB"
                 /clone="ZMMBTa0652L10"
                 /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                 cot selected genomic DNA library"
ORIGIN

Query Match   90.0%;   Score 18; DB 9;   Length 797;
Best Local Similarity 100.0%;   Pred. No. 0.051;
  Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   3 CAGTCGATATCGACTGAT   20
               ||||||||||||||||||
Db   366 CAGTCGATATCGACTGAT   349

RESULT 3
LOCUS   CG069001
DEFINITION   PUFNB43TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0687H14,
                genomic survey sequence.
ACCESSION   CG069001
VERSION   CG069001.1 GI:33941181
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 946)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
                Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
                Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
  COMMENT   Other_GSSs: PUFNB43TB
                Contact: Cathy Whitelaw
                TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..946
                 /organism="zea mays"
                 /mol_type="genomic DNA"
                 /strain="B73"
                 /db_xref="taxon:4577"
                 /clone_lib="ZM_0.6_1.0_KB"
                 /clone="ZMMBTa0687H14"
                 /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                 cot selected genomic DNA library"
ORIGIN

Query Match   90.0%;   Score 18; DB 9;   Length 946;
Best Local Similarity 100.0%;   Pred. No. 0.051;
  Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   3 CAGTCGATATCGACTGAT   20
               ||||||||||||||||||
Db   519 CAGTCGATATCGACTGAT   536

RESULT 4
LOCUS   CG069001/c
DEFINITION   PUFNB43TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0687H14,
                genomic survey sequence.
ACCESSION   CG069001
VERSION   CG069001.1 GI:33941181
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 946)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
                Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
                Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
  COMMENT   Other_GSSs: PUFNB43TB
                Contact: Cathy Whitelaw
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA
                Tel: 301-838-5843
                Fax: 301-838-0208
                Email: whitelaw@tigr.org
                Seq primer: TF
                Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..946
                 /organism="zea mays"
                 /mol_type="genomic DNA"
                 /strain="B73"
                 /db_xref="taxon:4577"
                 /clone_lib="ZM_0.6_1.0_KB"
                 /clone="ZMMBTa0687H14"
                 /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                 cot selected genomic DNA library"
ORIGIN

Query Match   90.0%;   Score 18; DB 9;   Length 946;
Best Local Similarity 100.0%;   Pred. No. 0.051;
  Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATCAGTCGATATCGACTG   18
               ||||||||||||||||||

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: October 27, 2004, 11:49:55 ; Search time 710.125 Seconds
(without alignments)
1331.870 Million cell updates/sec

Title: SEQ2-C2
Perfect score: 20
Sequence: 1 atcagtcgatatcgactgat 20
Scoring table: OLIGO_NUC
                  Gapop 60.0 , Gapext 60.0
Searched: 4526729 seqs, 2364849745 residues
Word size :   0
Total number of hits satisfying chosen parameters: 9053458
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :   GenEmbl:*
    1:   gb_ba:*
    2:   gb_htg:*
    3:   gb_in:*
    4:   gb_om:*
    5:   gb_ov:*
    6:   gb_pat:*
    7:   gb_ph:*
    8:   gb_pl:*
    9:   gb_pr:*
   10:   gb_ro:*
   11:   gb_rt:*
   12:   gb_sts:*
   13:   gb_sy:*
   14:   gb_un:*
   14:   gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	80.0	64915	5	AC091300	AC091300 Danio rer
2	16	80.0	64915	5	AC091300	AC091300 Danio rer
3	16	80.0	178225	5	BX005286	BX005286 Zebrafish
4	16	80.0	178225	5	BX005286	BX005286 Zebrafish
5	16	80.0	207611	2	BX004968	BX004968 Danio rer
6	16	80.0	207611	2	BX004968	BX004968 Danio rer
7	16	80.0	303250	1	AP001518	AP001518 Bacillus
8	16	80.0	303250	1	AP001518	AP001518 Bacillus
9	15	75.0	9179	1	AE014584	AE014584 Brucella
10	15	75.0	9179	1	AE014584	AE014584 Brucella
11	15	75.0	10362	1	AE011178	AE011178 Methanosa
12	15	75.0	10362	1	AE011178	AE011178 Methanosa
13	15	75.0	12244	1	AE009676	AE009676 Brucella
14	15	75.0	12244	1	AE009676	AE009676 Brucella
15	15	75.0	51601	1	AY374448	AY374448 Bacteriop
16	15	75.0	51601	7	AY374448	AY374448 Bacteriop
17	15	75.0	55896	5	BX276102	BX276102 Zebrafish
18	15	75.0	55896	5	BX276102	BX276102 Zebrafish
19	15	75.0	59820	2	BX649285_3	Continuation (4 of
20	15	75.0	59820	2	BX649285_3	Continuation (4 of
21	15	75.0	91609	5	BX323023	BX323023 Zebrafish
22	15	75.0	91609	5	BX323023	BX323023 Zebrafish
23	15	75.0	102704	5	AL672216	AL672216 Zebrafish
24	15	75.0	102704	5	AL672216	AL672216 Zebrafish
25	15	75.0	110590	5	BX001021	BX001021 Zebrafish
26	15	75.0	110590	5	BX001021	BX001021 Zebrafish
27	15	75.0	114157	5	BX294157	BX294157 Zebrafish
28	15	75.0	114157	5	BX294157	BX294157 Zebrafish
29	15	75.0	114924	2	AC149796	AC149796 Aedes aeg
30	15	75.0	114924	2	AC149796	AC149796 Aedes aeg
31	15	75.0	126202	5	AL591427	AL591427 Zebrafish
32	15	75.0	126202	5	AL591427	AL591427 Zebrafish
33	15	75.0	128296	5	BX897690	BX897690 Zebrafish
34	15	75.0	128296	5	BX897690	BX897690 Zebrafish
35	15	75.0	140982	5	BX842684	BX842684 Zebrafish
36	15	75.0	140982	5	BX842684	BX842684 Zebrafish
37	15	75.0	145496	2	CR545470	CR545470 Danio rer
38	15	75.0	145496	2	CR545470	CR545470 Danio rer
39	15	75.0	145777	2	BX511166	BX511166 Danio rer
40	15	75.0	145777	2	BX511166	BX511166 Danio rer
41	15	75.0	146052	2	CR384049	CR384049 Danio rer
42	15	75.0	146052	2	CR384049	CR384049 Danio rer
43	15	75.0	146575	2	BX957262	BX957262 Danio rer
44	15	75.0	146575	2	BX957262	BX957262 Danio rer
45	15	75.0	150351	2	BX247899	BX247899 Danio rer

ALIGNMENTS

RESULT 1
AC091300
LOCUS   AC091300   64915 bp   DNA   linear   VRT 16-OCT-2002
DEFINITION Danio rerio clone 127K3, complete sequence.
ACCESSION   AC091300
VERSION   AC091300.3 GI:24022405
KEYWORDS   HTG.
SOURCE   Danio rerio (zebrafish)
ORGANISM   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 64915)
AUTHORS   Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L.,
            Granite,S., Guan,X., Gupta,J., Ho,S.-L., Idol,J.R., Karlins,E.,
            Lee-Lin,S.-Q., Legaspi,R., Lim,M., Maduro,Q.L., Maduro,V.B.,
            Masiello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
            Pearson,R., Prasad,A., Shevchenko,Y., Snyder,B., Stantripop,S.,
            Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L.,
            Walker,M.A., Wetherby,K.D., Zhang,L.-H. and Green,E.D.
TITLE   NISC Comparative Sequencing Initiative
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 64915)
AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (13-APR-2001) NIH Intramural Sequencing Center,
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3 (bases 1 to 64915)
AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (07-JUL-2001) NIH Intramural Sequencing Center,
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   4 (bases 1 to 64915)
AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (16-OCT-2002) NIH Intramural Sequencing Center,
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT   On Oct 16, 2002 this sequence version replaced gi:14625272.
            --------- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC

Left column

Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
------- Project Information
Center: NISC
Center project name: aq1
Center clone name: 127K03

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

```
FEATURES             Location/Qualifiers
     source          1..64915
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="127K3"
                     /clone_lib="Incyte Genomics"
     misc_feature    26404..26418
                     /note="single clone coverage"
ORIGIN
```

```
RESULT 2
LOCUS       AC091300/c      64915 bp    DNA     linear   VRT 16-OCT-2002
DEFINITION  Danio rerio clone 127K3, complete sequence.
ACCESSION   AC091300
VERSION     AC091300.3  GI:24022405
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
```

REFERENCE   1 (bases 1 to 64915)
AUTHORS   Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L.,
            Granite,S., Guan,X., Gupta,J., Ho,S.-L., Idol,J.R., Karlins,E.,
            Lee-Lin,S.-Q., Legaspi,R., Lim,M., Maduro,Q.L.,
            Masiello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
            Pearson,R., Prasad,A., Shevchenko,Y., Snyder,B., Stantripop,S.,
            Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L.,
            Walker,M.A., Wetherby,K.D., Zhang,L.-H. and Green,E.D.
TITLE   NISC Comparative Sequencing Initiative
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 64915)
AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (13-APR-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3 (bases 1 to 64915)
AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (07-JUL-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   4 (bases 1 to 64915)

```
Query Match             80.0%;  Score 16;  DB 5;  Length 64915;
  Best Local Similarity 100.0%;  Pred. No. 12;
  Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        5 GTCGATATCGACTGAT 20
            ||||||||||||||||
Db    54116 GTCGATATCGACTGAT 54131
```

Right column

AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (16-OCT-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT   On Oct 16, 2002 this sequence version replaced gi:14626272.
            ------- Genome Center
            Center: NIH Intramural Sequencing Center
            Web site: http://www.nisc.nih.gov
            Contact: nisc_zoo@nhgri.nih.gov
            ------- Project Information
            Center project name: aq1
            Center clone name: 127K03

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

```
FEATURES             Location/Qualifiers
     source          1..64915
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="127K3"
                     /clone_lib="Incyte Genomics"
     misc_feature    26404..26418
                     /note="single clone coverage"
ORIGIN
```

```
RESULT 3
LOCUS       BX005286       178225 bp    DNA     linear   VRT 23-SEP-2003
DEFINITION  Zebrafish DNA sequence from clone DKEY-23719, complete sequence.
ACCESSION   BX005286
VERSION     BX005286.6  GI:33457028
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
```

REFERENCE   1 (bases 1 to 178225)
AUTHORS   Heath,P.
TITLE   Direct Submission
JOURNAL   Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Aug 5, 2003 this sequence version replaced gi:33412053.
            ------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk

            During sequence assembly data is compared from overlapping clones.

```
Query Match             80.0%;  Score 16;  DB 5;  Length 64915;
  Best Local Similarity 100.0%;  Pred. No. 12;
  Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        1 ATCAGTCGATATCGAC 16
            ||||||||||||||||
Db    54131 ATCAGTCGATATCGAC 54116
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: October 27, 2004, 11:49:55 ; Search time 710.125 Seconds
         (without alignments)
         1331.870 Million cell updates/sec

Title: SEQ2-2
Perfect score: 20
Sequence: 1 atcagtcgatcagtcgat 20
Scoring table: OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0 .
Searched: 4526729 seqs, 2364849745 residues
Word size :   0
Total number of hits satisfying chosen parameters: 9053458
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
   GenEmbl:*
    1: gb_ba:*
    2: gb_htg:*
    3: gb_in:*
    4: gb_om:*
    5: gb_ov:*
    6: gb_pat:*
    7: gb_ph:*
    8: gb_pl:*
    9: gb_pr:*
   10: gb_ro:*
   11: gb_sts:*
   12: gb_sy:*
   13: gb_un:*
   14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17	85.0	40701	3	CBRG21F09	AC084527 Caenorhab
	2	17	85.0	172064	9	AL445195	AL445195 Human DNA
	3	17	85.0	193858	3	AC016760	AC016760 Homo sapi
	4	16	80.0	15911	3	AF096709	AF096709 Drosophi
c	5	16	80.0	64915	5	AC091300	AC091300 Danio rer
c	6	16	80.0	207611	5	BX004968	BX004968 Danio rer
	7	15	75.0	4537	6	CO602864	CO602864 Sequence
	8	15	75.0	10052	3	CENVP13D12L	Al033535 Caenorhab
c	9	15	75.0	12761	1	AE005808	AE005808 Caulobact
c	10	15	75.0	5057	2	AC020176	AC020176 Drosophil
c	11	15	75.0	88249	2	BX897713	BX897713 Zebrafish
c	12	15	75.0	102986	2	AC020135	AC020135 Drosophil
c	13	15	75.0	110000	8	CR323137_11	Continuation (12 o
c	14	15	75.0	131395	9	AC084740	AC084740 Homo sapi
c	15	15	75.0	140310	2	AC016009	AC016009 Homo sapi
c	16	15	75.0	143287	2	AC010567	AC010567 Drosophil
c	17	15	75.0	149591	2	AC141821	AC141821 Drosophi
c	18	15	75.0	150830	5	BX470190	BX470190 Zebrafish
c	19	15	75.0	156799	2	AC079608	AC079608 Homo sapi
	20	15	75.0	157235	3	AC009338	AC009338 Drosophil
	21	15	75.0	157847	10	AC122511	AC122511 Mus muscu
	22	15	75.0	165137	8	CNS07YPU	AL713951 Oryza sat
	23	15	75.0	181056	3	AC009733	AC009733 Drosophil
	24	15	75.0	197278	9	AC007367	AC007367 Homo sapi
	25	15	75.0	199396	3	AC010064	AC010064 Drosophil
	26	15	75.0	225833	2	AC137266	AC137266 Rattus no
	27	15	75.0	226690	2	AC141328	AC141328 Mus muscu
	28	15	75.0	226904	3	AE003773	AE003773 Drosophil
	29	15	75.0	258149	2	AC098277	AC098277 Rattus no
	30	15	75.0	263169	2	BX927111	BX927111 Danio rer
	31	15	75.0	265197	2	AC113801	AC113801 Rattus no
	32	15	75.0	267667	2	AC094625	AC094625 Rattus no
c	33	15	75.0	294050	1	BX294142	BX294142 Pirellula
c	34	15	75.0	322554	3	AE003592	AE003592 Drosophil
	35	14	70.0	315	3	AY363977	AY363977 Xiphinema
	36	14	70.0	353	3	AY363976	AY363976 Xiphinema
	37	14	70.0	373	6	BD272941	BD272941 Materials
	38	14	70.0	373	6	BD272942	BD272942 Materials
c	39	14	70.0	373	6	BD273089	BD273089 Materials
c	40	14	70.0	446	6	CQ719271	CQ719271 Sequence
c	41	14	70.0	675	9	BD273088	BD273088 Materials
c	42	14	70.0	1344	6	BC032307	BC032307 Homo sapi
c	43	14	70.0	1377	6	AR393948	AR393948 Sequence
c	44	14	70.0	1377	6	AX366920	AX366920 Sequence
c	45	14	70.0	1403	1	AT72S1	X03933 Agrobacteri

ALIGNMENTS

RESULT 1
C3RG21F09   40701 bp   DNA   linear   INV 04-NOV-2000
LOCUS   CBRG21F09
DEFINITION Caenorhabditis briggsae cosmid G21F09, complete sequence.
ACCESSION   AC084527
VERSION   AC084527.1 GI:11094977
KEYWORDS   HTG.
SOURCE   Caenorhabditis briggsae
ORGANISM   Caenorhabditis briggsae
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1 (bases 1 to 40701)
AUTHORS   Washington University Genome Sequencing Center.
TITLE   The C. briggsae Genome Sequencing Project
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 40701)
AUTHORS   Waterston,R.
TITLE   Direct Submission
JOURNAL   Submitted (04-NOV-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT   Submitted by:
            Genome Sequencing Center
            Department of Genetics, Washington University,
            St. Louis, MO 63110, USA
            e-mail: jspieth@watson.wustl.edu

            NOTICE: This sequence may not be the entire insert of this clone.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we provide a small overlap between
            neighboring submissions.
FEATURES   Location/Qualifiers
     source   1..40701
                     /organism="Caenorhabditis briggsae"
                     /mol_type="genomic DNA"
                     /strain="Gujarat G16"
                     /db_xref="taxon:6238"
                     /clone="G21F09"
ORIGIN

Query Match   85.0%; Score 17; DB 3; Length 40701;
Best Local Similarity   100.0%; Pred. No. 4.2;

RESULT 2

Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   2 TCAGTCGATATCAGTCG 18
             ||||||||||||||||||
Db   5580 TCAGTCGATATCAGTCG 5596

LOCUS   AL445195   172064 bp   DNA   linear   PRI 15-DEC-2000
DEFINITION Human DNA sequence from clone RP11-427013 on chromosome 13,
            complete sequence.
ACCESSION   AL445195
VERSION   AL445195.4 GI:11878448
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 172064)
AUTHORS   Dunn,M.
TITLE   Direct Submission
JOURNAL   Submitted (15-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT   On Dec 18, 2000 this sequence version replaced gi:11545086.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated repeat sequence elements. Where the sequence is
            ambiguous, there is an annotation using the 'unsure' feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            RP11-427013 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
FEATURES   Location/Qualifiers
     source   1..172064
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone_lib="RPCI-11.2"
                     /clone="RP11-427013"
     repeat_region   283..443
                     /note="MLT1A2 repeat: matches 1..161 of consensus"
     repeat_region   457..520
                     /note="16 copies 4 mer ctct 85% conserved"
     repeat_region   527..721
                     /note="MLT1A2 repeat: matches 173..374 of consensus"
     repeat_region   1695..2057
                     /note="HAL1 repeat: matches 196..567 of consensus"
     repeat_region   2265..2478
                     /note="L2 repeat: matches 2534..2750 of consensus"
     repeat_region   2479..2598
                     /note="L2 repeat: matches 2534..2750 of consensus"
     repeat_region   2599..2979
                     /note="FLAM_C repeat: matches 1..131 of consensus"
     repeat_region   3012..3315
                     /note="L2 repeat: matches 1803..2099 of consensus"
     repeat_region   3391..3438
                     /note="12 copies 4 mer cata 75% conserved"
     repeat_region   5972..6119
                     /note="37 copies 4 mer aaat 58% conserved"
     repeat_region   6509..6643
                     /note="MER90 repeat: matches 5..136 of consensus"
     repeat_region   7417..7708
                     /note="AluSg repeat: matches 1..292 of consensus"
     repeat_region   7882..8514
                     /note="MSTA repeat: matches 1..426 of consensus"
     repeat_region   8771..8877
                     /note="L1PA7 repeat: matches 5514..6142 of consensus"
     repeat_region   11129..11365
                     /note="L1MA2 repeat: matches 6202..6308 of consensus"
     repeat_region   15199..15326
                     /note="MER46A repeat: matches 1..236 of consensus"
     repeat_region   15350..15755
                     /note="32 copies 4 mer atct 93% conserved"
     repeat_region   16332..16491
                     /note="MSTA repeat: matches 1..426 of consensus"
     repeat_region   17620..17936
                     /note="MIR repeat: matches 85..250 of consensus"
     repeat_region   18655..18816
                     /note="AluSg repeat: matches 1..308 of consensus"
     repeat_region   18973..19172
                     /note="MLT1J repeat: matches 312..472 of consensus"
     repeat_region   19680..20822
                     /note="MER30 repeat: matches 1..200 of consensus"
     repeat_region   20822..21420
                     /note="L1M2 repeat: matches -676..619 of consensus"
     repeat_region   21421..21714
                     /note="L1 repeat: matches 2331..2981 of consensus"
     repeat_region   21715..22215
                     /note="AluUb repeat: matches 1..301 of consensus"
     repeat_region   22194..22215
                     /note="L1 repeat: matches 2981..3465 of consensus"
     repeat_region   22216..22522
                     /note="AluSp repeat: matches 1..312 of consensus"
     repeat_region   22523..24006
                     /note="L1MA4A repeat: matches 4660..4681 of consensus"
     repeat_region   24192..24481
                     /note="L1MA4A repeat: matches 4681..6135 of consensus"
     repeat_region   24486..24641
                     /note="AluSg1 repeat: matches 14..301 of consensus"
     repeat_region   24705..24737
                     /note="L1MA4A repeat: matches 6125..6295 of consensus"
     repeat_region   24744..24997
                     /note="Tigger2a repeat: matches 412..444 of consensus"
     repeat_region   25057..25092
                     /note="Tigger2a repeat: matches 1..259 of consensus"
     repeat_region   25095..25398
                     /note="9 copies 4 mer tgtg 91% conserved"
     repeat_region   25708..26094
                     /note="AluJo repeat: matches 1..286 of consensus"
     repeat_region   26502..26914
                     /note="MLT1A2 repeat: matches 1..363 of consensus"
     repeat_region   26917..27308
                     /note="L2 repeat: matches 1964..2419 of consensus"
     repeat_region   28333..28467
                     /note="MSTA repeat: matches 1..425 of consensus"
     repeat_region   29041..29676
                     /note="FLAM_A repeat: matches 1..115 of consensus"
     repeat_region   29679..29841
                     /note="HSMAR2 repeat: matches 652..1298 of consensus"
     repeat_region   29845..29896
                     /note="FRAM repeat: matches 1..161 of consensus"
     repeat_region   29898..30015
                     /note="13 copies 4 mer aaat 76% conserved"
     repeat_region   30046..30270
                     /note="FLAM_C repeat: matches 1..118 of consensus"
     repeat_region   30271..30622
                     /note="HSMAR2 repeat: matches 422..644 of consensus"

repeat_region   /note="THE1C repeat: matches 1. .371 of consensus"
30623. .30794
repeat_region   /note="HSMAR2 repeat: matches 252. .422 of consensus"
30795. .31170
repeat_region   /note="THE1C repeat: matches 1. .371 of consensus"
31171. .31411
repeat_region   /note="HSMAR2 repeat: matches 2. .252 of consensus"
34651. .35102
repeat_region   /note="MER4C repeat: matches 1. .465 of consensus"
35216. .35290
repeat_region   /note="MER53 repeat: matches 7. .82 of consensus"
35291. .35958
repeat_region   /note="MER4A repeat: matches 1. .465 of consensus"
35965. .36068
repeat_region   /note="MER53 repeat: matches 1. .113 of consensus"
36798. .37106
repeat_region   /note="AluY repeat: matches 3. .309 of consensus"
37396. .37443
repeat_region   /note="24 copies 2 mer at 79% conserved"
41913. .42368
repeat_region   /note="12 repeat: matches 2270. .2749 of consensus"
44620. .44667
repeat_region   /note="24 copies 2 mer tt 72% conserved"
45999. .46780
repeat_region   /note="LTR17 repeat: matches 1. .780 of consensus"
46970. .47275
repeat_region   /note="AluSx repeat: matches 1. .305 of consensus"
47305. .47695
repeat_region   /note="LTR7 repeat: matches 1. .449 of consensus"
48464. .48515
repeat_region   /note="26 copies 2 mer ca 86% conserved"
48473. .48516
repeat_region   /note="11 copies 4 mer acac 90% conserved"
48570. .48757
repeat_region   /note="12 repeat: matches 2114. .2303 of consensus"
48774. .49037
repeat_region   /note="MER7A repeat: matches 68. .336 of consensus"
49073. .49204
repeat_region   /note="33 copies 4 mer atta 59% conserved"
49252. .49331
consensus"   /note="Tigger3(Golem) repeat: matches 1. .2772 of
49457. .49778
repeat_region   /note="11PB2 repeat: matches 5815. .6152 of consensus"
50893. .50966
repeat_region   /note="37 copies 2 mer aa 64% conserved"
53081. .54131
repeat_region   /note="12 repeat: matches 33. .1147 of consensus"
54270. .54582
repeat_region   /note="AluSx repeat: matches 1. .309 of consensus"
54919. .55259
repeat_region   /note="THE1C repeat: matches 1. .371 of consensus"
59089. .59382
repeat_region   /note="MIT1A1 repeat: matches 10. .365 of consensus"
60033. .60174
repeat_region   /note="21 copies 2 mer ca 90% conserved"
60141. .60176
repeat_region   /note="9 copies 4 mer caca 94% conserved"
60574. .60918
repeat_region   /note="MIT2PB repeat: matches 1. .344 of consensus"
60963. .60988
repeat_region   /note="13 copies 2 mer ta 92% conserved"
60989. .61041

Query Match   85.0%;   Score 17;   DB 9;   Length 172064;
Best Local Similarity 100.0%;   Pred. No. 4.7;
Matches 17;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   3 CAGTCGGATATCAGTCGA 19
            |||||||||||||||||
Db   169291 CAGTCGGATATCAGTCGA 169307

RESULT 3
LOCUS   AC016760   193858 bp   DNA   linear   HTG 07-JUL-2000
DEFINITION   Homo sapiens chromosome 13 clone RP11-536M12, WORKING DRAFT
               SEQUENCE, 12 unordered pieces.
ACCESSION   AC016760
VERSION   AC016760.4 GI:8568907
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 193858)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Homo sapiens clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 193858)
AUTHORS   Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (04-DEC-1999) Genome Sequencing Center, Washington
               University School of Medicine, 4444 Forest Park Parkway, St. Louis,
               MO 63108, USA
COMMENT   On Jun 16, 2000 this sequence version replaced gi:6850523.

               --------------- Genome Center -----------------
               Center: Washington University Genome Sequencing Center
               Center code: WUGSC
               Web site: http://genome.wustl.edu/gsc/index.shtml
               --------------- Project Information -----------------
               Center project name: H_NH0536M12
               --------------- Summary Statistics -----------------
               Sequencing vector: M13; 88%
               Sequencing vector: plasmid; 12%
               Chemistry: Dye-primer ET; 84% of reads
               Chemistry: Dye-terminator Big Dye; 16% of reads
               Assembly program: Phrap; version 0.990319
               Consensus quality: 181257 bases at least Q40
               Consensus quality: 185575 bases at least Q30
               Consensus quality: 188576 bases at least Q20
               Insert size: 184000; agarose-fp
               Insert size: 192758; sum-of-contigs
               Quality coverage: 4.34 in Q20 bases; agarose-fp
               Quality coverage: 4.17 in Q20 bases; sum-of-contigs
               ----
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 12 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *   1   1629: contig of 1629 bp in length
             *   1630   1729: gap of unknown length
             *   1730   4727: contig of 2998 bp in length
             *   4728   4827: gap of unknown length
             *   4828   8897: contig of 4070 bp in length
             *   8898   8997: gap of unknown length
             *   8998   14918: contig of 5921 bp in length
             *   14919   15018: gap of unknown length
             *   15019   21728: contig of 6710 bp in length
             *   21729   21828: gap of unknown length
             *   21829   36092: contig of 14264 bp in length
             *   36093   36192: gap of unknown length
             *   36193   51110: contig of 14918 bp in length
             *   51111   51210: gap of unknown length
             *   51211   71759: contig of 20549 bp in length
             *   71760   71859: gap of unknown length
             *   71860   97429: contig of 25570 bp in length
             *   97430   97529: gap of unknown length
             *   97530   122384: contig of 24855 bp in length
             *   122385   122484: gap of unknown length
             *   122485   154483: contig of 31999 bp in length

```
          *   154484   154583: gap of unknown length
          *   154584   193858: contig of 39275 bp in length.
                  1. .193858
FEATURES          Location/Qualifiers
source            1. .193858
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="13"
                  /clone="RP11-536M12"
misc_feature      1. .1629
                  /note="assembly_name:Contig7"
misc_feature      1730. .4727
                  /note="assembly_name:Contig8"
misc_feature      4828. .8897
                  /note="assembly_name:Contig9"
misc_feature      8998. .14918
                  /note="assembly_name:Contig10
                  clone_end:T7
                  vector_side:right"
misc_feature      15019. .21728
                  /note="assembly_name:Contig11"
misc_feature      21829. .36092
                  /note="assembly_name:Contig12"
misc_feature      36193. .51110
                  /note="assembly_name:Contig13"
misc_feature      51211. .71759
                  /note="assembly_name:Contig14"
misc_feature      71860. .97429
                  /note="assembly_name:Contig15"
misc_feature      97530. .122384
                  /note="assembly_name:Contig16"
misc_feature      122485. .154483
                  /note="assembly_name:Contig17"
misc_feature      154584. .193858
                  /note="assembly_name:Contig18
                  clone_end:SP6
                  vector_side:right"
ORIGIN
```

Query Match   85.0%; Score 17; DB 2; Length 193858;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3 CAGTCGATATCAGTCGA 19
           |||||||||||||||||
Db  180196 CAGTCGATATCAGTCGA 180212
```

RESULT 4

```
LOCUS       AF096709/c      15911 bp    DNA     linear   INV 09-JAN-1999
DEFINITION  Drosophila virilis tyrosyl-tRNA synthetase gene, partial cds; and
            failed axon connections protein (fax) gene, complete cds.
ACCESSION   AF096709
VERSION     AF096709.1  GI:4128222
KEYWORDS
SOURCE      Drosophila virilis
ORGANISM    Drosophila virilis
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 15911)
AUTHORS     Bedian,V., Jungklaus,C.E., Cardoza,L. and von Kalm,L.
TITLE       Kinase activity and genetic characterization of a growth related
            antigen of Drosophila
JOURNAL     Dev. Genet. 12 (3), 188-195 (1991)
MEDLINE     91330447
PUBMED      1686624
REFERENCE   2  (bases 1 to 15911)
AUTHORS     Hill,K.K., Bedian,V., Juang,J.L. and Hoffmann,F.M.
TITLE       Genetic interactions between the Drosophila Abelson (Abl) tyrosine
            kinase and failed axon connections (fax), a novel protein in axon
            bundles
JOURNAL     Genetics 141 (2), 595-606 (1995)
MEDLINE     96109606
PUBMED      8647396
REFERENCE   3  (bases 1 to 15911)
AUTHORS     Geiger,B.A., Park,C., Rees,H., Frederick,J. and Bedian,V.
TITLE       Direct Submission
JOURNAL     Submitted (02-OCT-1998) Genetics, University of Pennsylvania, 476
            CRB, 415 Curie Blvd, Philadelphia, PA 19104, USA
FEATURES    Location/Qualifiers
source      1. .15911
            /organism="Drosophila virilis"
            /mol_type="genomic DNA"
            /db_xref="taxon:7244"
            /chromosome="3"
            /map="72F"
            /note="putative chromosome and map positions based on
            Drosophila melanogaster map positions"
mRNA        complement(join(<220. .285,341. .1021))
            /product="tyr-tRNA synthetase"
CDS         complement(join(<220. .285,341. .986))
            /codon_start=1
            /product="tyrosyl-tRNA synthetase"
            /protein_id="AAD03560.1"
            /db_xref="GI:4128224"
            /translation="MVGTTLSADEKKQLITRNLQ3TLGEDKLNKVLQERDLKIYWGTA
            TGKPHVAYFVPMSKIADFLKAGCEVTILFADLHAYLDNMKAPWPLLELRTKYYEQVI
            KAMLSSIGVPLDKLKFVKGSSDYQLSREYTLDVYKLSSLVTHDAKXAGAFVVKQVEHP
            IISCLLYPGLQGIDEEYLKVDAQFGGVDQRKIFTFSEKYLPQLGYSKIHFMNPVPG
            LAGGKMSSSEEDSKIDLLD"
gene        <6683. .>14015
            /gene="fax"
mRNA        join(<6683. .7120,11368. .11706,13319. .13662,13728. .13830,
            13962. .14015)
            /gene="fax"
            /product="failed axon connections protein"
CDS         join(6683. .7120,11368. .11706,13319. .13662,13728. .13830,
            13962. .14015)
            /gene="fax"
            /codon_start=1
            /product="failed axon connections protein"
            /protein_id="AAD03559.1"
            /db_xref="GI:4128223"
            /translation="MATEVAQIPABETPAVAAAAAATTENAAAAEQA3KPAAEPVAVA
            AVAENAAANNAQDNNKEDAKASPAAAADGAAAADAKDGDGGVAAAPAKSEAPAAQKEN
            VHKTNFEKDIIYLYQFSRSPLLPSLSPYCLKVETWLRLVALKYENVHKMRFRSKKGQ
            LPPIELNGEEIADSAIIIKELSAKYDKNLDAGLTAEQRNVSYATIAMIENHLIWIFY
            WRAKYPDNVLKGYKVNLQEHLGLRLPNSILNFFFKITPGRKWFQGTKKLKAHGICVHS
            ABBIEEFGKNDLKVLSEMLDCKPFFGDETTLDVVAFAVLSQLHYLSKDIAYPLRDY
            MTEKCPNIIGHVSRMKDKCFPDWDEICTKLDLNAHIPKPEPETKEGKSGGGEQEKSNEO
            EGTDGDKIEKELEKDKSNEKESTEENKEKEETK"
ORIGIN
```

Query Match   80.0%; Score 16; DB 3; Length 15911;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       5 GTCGATATCAGTCGAT 20
           ||||||||||||||||
Db    8933 GTCGATATCAGTCGAT 8918
```

RESULT 5

```
LOCUS       AC091300        64915 bp    DNA     linear   VRT 16-OCT-2002
DEFINITION  Danio rerio clone 127K3, complete sequence.
ACCESSION   AC091300
VERSION     AC091300.3  GI:24022405
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Danio.
            Cypriniformes; Cyprinidae; Danio.
```